

# SEARCH REQUEST FORM

6-214

Requestor's Name: Sheela Huff Serial Number: 08/231565  
Date: 6/9/97 Phone: 305-7866 Art Unit: 1806

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

PLEASE SEARCH SEQ. ID # 4, 17, 18,  
31-39.

TXS-  
Sheela

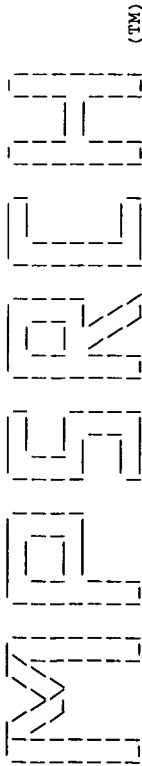
## STAFF USE ONLY

Date completed: 6-10-97  
Searcher: BOB x4291  
Terminal time: 4  
Elapsed time: prep 15  
CPU time: \_\_\_\_\_  
Total time: \_\_\_\_\_  
Number of Searches: \_\_\_\_\_  
Number of Databases: 3

Search Site  
\_\_\_\_ STIC  
\_\_\_\_ CM-1  
\_\_\_\_ Pre-S  
Type of Search  
\_\_\_\_ N.A. Sequence  
12 A.A. Sequence  
\_\_\_\_ Structure  
\_\_\_\_ Bibliographic

Vendors  
\_\_\_\_ IG Suite  
\_\_\_\_ STN  
\_\_\_\_ Dialog  
\_\_\_\_ APS  
\_\_\_\_ Geninfo  
\_\_\_\_ SDC  
\_\_\_\_ DARC/Questel  
MP Other

\*\*\*\*\*


 The logo for Miles Research, featuring the word "Miles" in a large, stylized, outlined font, with "Research" in a smaller, outlined font below it. To the right of the word "Research" is a small trademark symbol (TM).

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:18:04 1997; MasPar time 1.85 Seconds

Tabular output not generated. 52.885 Million cell updates/sec

Title: >US-08-231-565A-4  
 Description: (1-9) from US08231565A.pep  
 Perfect Score: 56  
 Sequence: 1 AAGIGILTV 9

Scoring table: PAM 150  
 Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-geneseq26  
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
 14:part14 15:part15 16:part16 17:part17 18:part18  
 19:part19

Statistics: Mean 15.093; Variance 44.646; scale 0.338

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	56	100.0	9	15	MART-1 melanoma anti	4.29e+00
2	56	100.0	10	15	MART-1 melanoma anti	4.29e+00
3	56	100.0	10	15	MART-1 melanoma anti	4.29e+00
4	56	100.0	118	12	Tumour rejection anti	4.29e+00
5	56	100.0	118	15	MART-1 melanoma anti	4.29e+00
6	52	92.9	9	15	Modified MART-1 melan	1.25e+01
7	52	92.9	9	15	Modified MART-1 melan	1.25e+01
8	52	92.9	9	15	Modified MART-1 melan	1.25e+01
9	51	91.1	9	15	Modified MART-1 melan	1.62e+01
10	50	89.3	9	15	Modified MART-1 melan	2.11e+01
11	49	87.5	9	15	Modified MART-1 melan	2.74e+01
12	48	85.7	9	15	Modified MART-1 melan	3.55e+01
13	48	85.7	9	15	MART-1 melanoma anti	3.55e+01
14	48	85.7	9	15	Modified MART-1 melan	3.55e+01
15	48	85.7	9	15	Modified MART-1 melan	3.55e+01
16	48	85.7	9	15	Modified MART-1 melan	3.55e+01
17	48	85.7	9	15	Modified MART-1 melan	3.55e+01
18	48	85.7	9	15	Modified MART-1 melan	3.55e+01
19	47	83.9	479	3	Sequence encoded by t	4.59e+01
20	47	83.9	479	1	Complete sequence of	4.59e+01

21	43	76.8	9	15	R84790	Modified MART-1 melan	1.27e+02
22	42	75.0	231	15	R28440	Cephalosporin C acety	1.62e+02
23	41	73.2	9	15	R84791	Modified MART-1 melan	2.08e+02
24	41	73.2	9	15	R84798	Modified MART-1 melan	2.08e+02
25	41	73.2	9	15	R84789	Modified MART-1 melan	2.08e+02
26	41	73.2	9	15	R84800	Modified MART-1 melan	2.08e+02
27	41	73.2	9	15	R84799	Modified MART-1 melan	2.08e+02
28	41	73.2	25	8	R39763	Melittin delta 4 (del	2.08e+02
29	41	73.2	1464	18	R80970	Human excitatory amin	2.08e+02
30	41	73.2	1464	18	R86039	Human N-methyl-D-aspa	2.08e+02
31	41	73.2	1464	10	R55529	Human NMDA R2A recept	2.08e+02
32	40	71.4	1087	17	R96028	P. gingivalis haemagg	2.66e+02
33	40	71.4	1358	17	R96032	P. gingivalis haemagg	2.66e+02
34	40	71.4	1687	17	R96033	P. gingivalis haemagg	2.66e+02
35	40	71.4	1704	13	R70188	Arg-gingipain-2 prepo	2.66e+02
36	40	71.4	1732	17	R96029	P. gingivalis porphyp	2.66e+02
37	39	69.6	227	16	R86834	Human CD9 sequence.	3.39e+02
38	39	69.6	228	5	R27525	Metastasis controllin	3.39e+02
39	39	69.6	516	16	R94188	Pea ADPG-Phase B72A s	3.39e+02
40	39	69.6	813	12	R62754	Sefc sequence.	3.39e+02
41	39	69.6	2227	3	P60066	Sequence of viral L43	3.39e+02
42	39	69.6	2227	1	R05697	Attenuated hepatitis	3.39e+02
43	38	67.9	399	19	W04215	Rat vas deferens P2x	4.31e+02
44	38	67.9	434	18	R97745	Serine hydroxymethyl	4.31e+02
45	38	67.9	685	18	W04209	Mutant streptavidin/1	4.31e+02

## ALIGNMENTS

RESULT 1  
 ID R84196 standard; Peptide; 9 AA.  
 AC R84196;  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen immunogenic peptide M9-2.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine  
 QS Synthetic.  
 PN W09529193-A2.  
 PD 02-NOV-1995  
 PE 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host-cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 12; Page 117; 184pp; English.  
 CC Immunogenic Peptide M9-2 is based on the melanoma antigen (MART-1)  
 CC (see R84212). M9-2 may be modified to improve immunogenicity  
 CC (see R84783-R84800) and used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 100.0%; Score 56; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.29e+00;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aagigiltv 9  
 |||||  
 QY 1 AAGIGILTV 9

RESULT 2  
 ID R84198 standard; Peptide; 10 AA.  
 AC R84198;  
 DT 20-APR-1996 (first entry)

DE MART-1 melanoma antigen immunogenic peptide M10-4.  
 KW MART-1; M10-4; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 12; Page 122; 184pp; English.  
 CC Immunogenic peptide M10-4 is a derivative of peptide M9-2 (R84196)  
 CC which is based on the melanoma antigen (MART-1) (see R84212).  
 CC M9-2 may be modified to improve immunogenicity (see R84783-R84800)  
 CC and used in medicaments for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against MART-1 and its  
 CC immunogenic peptides may be used in the detection and isolation of  
 CC MART-1 from a sample, the detection of which is indicative of a  
 CC disease state (melanoma or metastatic melanoma).  
 CC See also R84196.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 4.29e+00;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aagigiltv 9  
 |||||  
 QY 1 AAGIGILTV 9

RESULT 3  
 ID R84197 standard; Peptide; 10 AA.  
 AC R84197;  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen immunogenic peptide M10-3 derivative.  
 KW MART-1; M10-3; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 12; Page 122; 184pp; English.  
 CC Immunogenic peptide M10-3 is a derivative of peptide M9-2 (R84196)  
 CC which is based on the melanoma antigen (MART-1) (see R84212).  
 CC M9-2 may be modified to improve immunogenicity (see R84783-R84800)  
 CC and used in medicaments for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against MART-1 and its  
 CC immunogenic peptides may be used in the detection and isolation of  
 CC MART-1 from a sample, the detection of which is indicative of a  
 CC disease state (melanoma or metastatic melanoma).  
 CC See also R84198  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 56; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 4.29e+00;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 2 aagigiltv 10  
 |||||  
 QY 1 AAGIGILTV 9

RESULT 4  
 ID R63158 standard; Protein; 118 AA.  
 AC R63158;  
 DT 26-MAY-1995 (first entry)  
 DE Tumour rejection antigen precursor.  
 KW Tumour rejection antigen; precursor; HLA-A2 molecule; tyrosinase;  
 KW isolation; melanoma; cell line; LB-39-MEL; diagnosis; vaccine;  
 KW therapy.  
 OS Homo sapiens  
 PN WO9529193-A2.  
 PD 29-SEP-1994.  
 PF 09-MAR-1994; U02487.  
 PR 18-MAR-1993; US-032978.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PI Boon-Fallour T, Brichard V, De Plaen E, Traversari C;  
 PI Van Pel A, Wolfel T;  
 DR WPI; 94-316544/39.  
 DR N-PSDB; 076370.  
 PT Nucleic acid coding for a tumour rejection antigen precursor - is  
 PT used for developing prods. for diagnosis or treatment of expression  
 PT related disorders, partic. melanoma  
 PS Claim 5; Page 14; 26pp; English.  
 CC This sequence represents the tumour rejection antigen precursor which is  
 CC processed to a tumour rejection antigen presented by HLA-A2 molecules.  
 CC The tumour rejection antigen is not related to tyrosinase. The CDNA  
 CC encoding this sequence was isolated from the melanoma cell line,  
 CC LB-39-MEL. The tumour rejection antigen may be used for diagnosis or  
 CC in vaccines or for therapy of disorders characterised by the expression  
 CC of the tumour rejection antigen precursor, particularly melanoma.  
 SQ Sequence 118 AA;

Query Match 100.0%; Score 56; DB 12; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 4.29e+00;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 aagigiltv 35  
 |||||  
 QY 1 AAGIGILTV 9

RESULT 5  
 ID R84212 standard; Protein; 118 AA.  
 AC R84212;  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen.  
 KW MART-1; melanoma antigen recognised by T-cell; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian.  
 FH Key Location/Qualifiers  
 FT Region 27..47  
 FT /note= "hydrophobic region"  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DR N-PSDB; T02714.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 11; Page 117; 184pp; English.  
 CC The melanoma antigen (MART-1) is produced by recombinant DNA

CC methods, i.e. preferably using a baculovirus vector for expression  
 CC in insect cell cultures. MART-1 protein is a source of immunogenic  
 CC peptides (see R84196 for peptide M9-2) which are optionally modified  
 CC (see R84783-R84800) and used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 118 AA;

Query Match 100.0%; Score 56; DB 15; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 4.29e+00;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 aagigiltv 35  
 |||||

Qy 1 AAGIGILTV 9

RESULT 6  
 ID R84788 standard; Peptide; 9 AA.  
 AC R84788;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1Y.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 25-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 92.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.25e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||

Qy 2 AGIGILTV 9

RESULT 7  
 ID R84787 standard; Peptide; 9 AA.  
 AC R84787;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1F.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 25-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 92.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.25e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||

Qy 2 AGIGILTV 9

RESULT 7  
 ID R84787 standard; Peptide; 9 AA.  
 AC R84787;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1F.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 25-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 92.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.25e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||

Qy 2 AGIGILTV 9

RESULT 9  
 ID R84785 standard; Peptide; 9 AA.  
 AC R84785;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2I.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;

PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 92.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.25e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||

Qy 2 AGIGILTV 9

RESULT 8  
 ID R84786 standard; Peptide; 9 AA.  
 AC R84786;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1W.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 25-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 92.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.25e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||

Qy 2 AGIGILTV 9

RESULT 9  
 ID R84785 standard; Peptide; 9 AA.  
 AC R84785;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2I.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;



```

RESULT 11
IID R84783 standard; Peptide; 9 AA.
AC R84783;
DT 25-APR-1996 (first entry)
DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2L.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PPR 22-APR-1994; US-231565.
PPR 05-APR-1995; US-417174.
PPA (US$H ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
PI WPI; 95-382963/49.
DR DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 105; 184pp; English.
PS R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is
CCC an immunogenic peptide based on the melanoma antigen (MART-1) (see
CCC R84212). The peptides are used in medicaments for the treatment or
CCC prevention (by immunization) of melanoma. Antibodies against MART-1
CCC and its immunogenic peptides may be used in the detection and
CCC isolation of MART-1 from a sample, the detection of which is
CCC indicative of a disease state (melanoma or metastatic melanoma).
CCC Sequence 9 AA;
SQ

Query Match 87.5%; Score 49; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 2.74e+01;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 aligilgtv 9
| | | | | | |
QY 1 AAGIGILTV 9

RESULT 12
IID R84792 standard; Peptide; 9 AA.
AC R84792;
DT 25-APR-1996 (first entry)
DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1K2L.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PPR 22-APR-1994; US-231565.
PPR 05-APR-1995; US-417174.
PPA (US$H ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
PI WPI; 95-382963/49.
DR DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 105; 184pp; English.
PS R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is
CCC an immunogenic peptide based on the melanoma antigen (MART-1) (see
CCC R84212). The peptides are used in medicaments for the treatment or
CCC prevention (by immunization) of melanoma. Antibodies against MART-1
CCC and its immunogenic peptides may be used in the detection and
CCC isolation of MART-1 from a sample, the detection of which is
CCC indicative of a disease state (melanoma or metastatic melanoma).
CCC Sequence 9 AA;
SQ

Query Match 85.7%; Score 48; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.55e+01;

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigiltv 9  
|||||||

QY 3 GIGILTV 9

# RESULT 13

ID R84764 standard; Peptide; 9 AA.

AC R84764; (first entry)

DE MART-1 melanoma antigen immunogenic peptide M9-3.

KW MART-1; M9-3; melanoma antigen recognised by T-cells; melanoma;

KW metastatic melanoma; tumour-associated antigen;

KW immunogenic peptide; diagnosis; prognosis; prophylaxis;

KW therapy; vaccine.

OS Synthetic.

PN WO9529193-A2.

PD 02-NOV-1995.

PF 21-APR-1995; U05063.

PR 22-APR-1994; US-231565.

PR 05-APR-1995; US-417174.

PA (USSH ) US SEC DEPT HEALTH.

PI Kawakami Y, Rosenberg SA;

DR WPI; 95-382963/49.

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also

PT vectors, host cells and antibodies, used to detect, treat and

PT immunise animal against melanoma.

PS Example 2; Page 72; 184pp; English.

CC Immunogenic peptide M9-3 is based on the melanoma antigen (MART-1)

CC (see R84212). M9-3 may be modified to improve immunogenicity

CC (see R84763-R84782) and used in medicaments (vaccines) for

CC treatment or prevention (by immunization) of melanoma.

CC Antibodies against MART-1 and its immunogenic peptides may be used

CC in the detection and isolation of MART-1 from a sample, the

CC detection of which is indicative of a disease state (melanoma

CC or metastatic melanoma).

CC See also R84196-R84198.

SQ Sequence 9 AA;

# Query Match

Best Local Similarity 85.7%; Score 48; DB 15; Length 9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 gigiltv 7

|||||||

QY 3 GIGILTV 9

# RESULT 14

ID R84796 standard; Peptide; 9 AA.

AC R84796; (first entry)

DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1r2L.

KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;

KW metastatic melanoma; tumour-associated antigen;

KW immunogenic peptide; diagnosis; prognosis; prophylaxis;

KW therapy; vaccine.

OS Synthetic.

PN WO9529193-A2.

PD 02-NOV-1995.

PF 21-APR-1995; U05063.

PR 22-APR-1994; US-231565.

PR 05-APR-1995; US-417174.

PA (USSH ) US SEC DEPT HEALTH.

PI Kawakami Y, Rosenberg SA;

DR WPI; 95-382963/49.

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also

PT vectors, host cells and antibodies, used to detect, treat and

PT immunise animal against melanoma.

PS Example 5; Page 105; 184pp; English.

CC R84783-800 are M9-2 peptides modified to improve immunogenicity, M9-2 is

CC an immunogenic peptide based on the melanoma antigen (MART-1) (see

CC R84212). The peptides are used in medicaments for the treatment or  
prevention (by immunization) of melanoma. Antibodies against MART-1  
and its immunogenic peptides may be used in the detection and  
isolation of MART-1 from a sample, the detection of which is  
indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 85.7%; Score 48; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.55e+01;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigiltv 9

|||||||

QY 3 GIGILTV 9

# RESULT 15

ID R84795 standard; Peptide; 9 AA.

AC R84795;

DT 25-APR-1996 (first entry)

DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1W2L.

KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;

KW metastatic melanoma; tumour-associated antigen;

KW immunogenic peptide; diagnosis; prognosis; prophylaxis;

KW therapy; vaccine.

OS Synthetic.

PN WO9529193-A2.

PD 02-NOV-1995.

PF 21-APR-1995; U05063.

PR 22-APR-1994; US-231565.

PR 05-APR-1995; US-417174.

PA (USSH ) US SEC DEPT HEALTH.

PI Kawakami Y, Rosenberg SA;

DR WPI; 95-382963/49.

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also

PT vectors, host cells and antibodies, used to detect, treat and

PT immunise animal against melanoma.

PS Example 5; Page 105; 184pp; English.

CC R84783-800 are M9-2 peptides modified to improve immunogenicity, M9-2 is

CC an immunogenic peptide based on the melanoma antigen (MART-1) (see

CC R84212). The peptides are used in medicaments for the treatment or

CC prevention (by immunization) of melanoma. Antibodies against MART-1

CC and its immunogenic peptides may be used in the detection and

CC isolation of MART-1 from a sample, the detection of which is

CC indicative of a disease state (melanoma or metastatic melanoma).

SQ Sequence 9 AA;

# Query Match

Best Local Similarity 85.7%; Score 48; DB 15; Length 9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigiltv 9

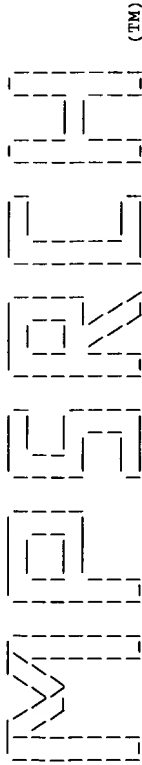
|||||||

QY 3 GIGILTV 9

Search completed: Tue Jun 10 11:18:13 1997

Job time : 9 secs.

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit,  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:17:34 1997; MasPar time 5.06 Seconds  
Tabular output not generated. 50.701 Million cell updates/sec

Title: >US-08-231-565A-4.  
Description: (1-9) from US08231565A.pap  
Perfect Score: 56  
Sequence: 1 AAGIGILTV 9

Scoring table: PAM 150  
Gap 15

Searched: 89912 seqs, 28507787 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir50  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 21.480; Variance 27.992; scale 0.767

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	56	100.0	118 13	A55253 melanoma antigen MAR	1.17e+01
2	56	100.0	118 13	I38506 melan-A protein - hu	1.17e+01
3	47	83.9	101 10	A33351 H+-transporting ATP	8.60e+00
4	47	83.9	479 4	VGBEPB glycoprotein gIII pr	8.60e+00
5	46	82.1	231 12	S48276 YSA1 protein - yeast	1.35e+01
6	45	80.4	339 16	S62369 methylobalamine: Coe	2.10e+01
7	45	80.4	420 14	S59131 Kan-1 protein - rat	2.10e+01
8	44	78.6	635 10	S19011 endo-1,4-beta-xylosa	3.24e+01
9	43	76.8	132 5	S01903 H+-transporting ATP	4.97e+01
10	43	76.8	493 3	ACMSE nicotinic acetylchol	4.97e+01
11	43	76.8	611 8	JT0592 hypothetical protein	4.97e+01
12	42	75.0	19 10	A34624 31K flagellin - Meth	7.57e+01
13	42	75.0	110 4	VHBPDL major capsid protein	7.57e+01
14	42	75.0	216 10	C41316 flagellin B2 precurs	7.57e+01
15	42	75.0	218 10	B41316 flagellin B1 precurs	7.57e+01
16	42	75.0	308 9	H64153 hypothetical protein	7.57e+01
17	42	75.0	333 1	DEALX3 aceoic[2,6-dichloro	7.57e+01
18	42	75.0	461 9	G64107 thiophene and furan	7.57e+01
19	42	75.0	501 10	S18573 L-lysine transport p	7.57e+01
20	42	75.0	530 11	S61034 AIP2 protein - yeast	7.57e+01
21	42	75.0	1530 11	S52239 hba2 protein - fissi	7.57e+01

22	41	73.2	226 3	JX0221	CD9 antihetigen - bovine	1.14e+02
23	41	73.2	271 11	S38115	hypothetical protein	1.14e+02
24	41	73.2	323 9	A64054	selenium metabolism	1.14e+02
25	41	73.2	345 1	KZEC	alkaline phosphatase	1.14e+02
26	41	73.2	401 16	S62465	hypothetical protein	1.14e+02
27	41	73.2	753 1	A42863	5-methyltetrahydropt	1.14e+02
28	41	73.2	845 14	A40016	matrin 3 - rat	1.14e+02
29	41	73.2	1436 11	S67655	probable membrane pr	1.14e+02
30	41	73.2	1484 13	S47555	N-methyl-D-aspartate	1.14e+02
31	40	71.4	78 7	S11923	gas-vesicle protein	1.72e+02
32	40	71.4	79 7	S06787	hypothetical protein	1.72e+02
33	40	71.4	79 7	S01422	gas-vesicle protein	1.72e+02
34	40	71.4	219 10	A41316	flagellin A precursor	1.72e+02
35	40	71.4	264 9	PS0398	hypothetical protein	1.72e+02
36	40	71.4	461 9	S01840	nitrogenase (EC 1.18	1.72e+02
37	40	71.4	491 9	I40990	signal transducer am	1.72e+02
38	40	71.4	491 9	I40991	signal transducer am	1.72e+02
39	40	71.4	491 9	I40989	signal transducer am	1.72e+02
40	40	71.4	566 9	A46162	nitrate sensor prote	1.72e+02
41	40	71.4	626 11	B53318	malate dehydrogenase	1.72e+02
42	40	71.4	793 12	S65240	probable membrane pr	1.72e+02
43	40	71.4	1526 9	S49763	protease precursor -	1.72e+02
44	40	71.4	1704 9	A55426	gingipain - Porphyro	1.72e+02
45	40	71.4	3519 10	S43048	polyketide synthase	1.72e+02

ALIGNMENTS

RESULT 1  
ENTRY A55253 #type complete  
TITLE melanoma antigen MART-1 - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change  
06-Feb-1995  
ACCESSIONS A55253  
REFERENCE A55253  
#authors Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.;  
Rivoltini, L.; Topalian, S.L.; Miki, T.; Rosenberg, S.A.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1994) 91:3515-3519  
#title Cloning of the gene coding for a shared human melanoma  
antigen recognized by autologous T cells infiltrating into  
tumor.  
#accession A55253  
##status preliminary  
##molecule\_type mRNA  
##residues 1-118 #label KAW  
##cross-references GB:006452  
SUMMARY #length 118 #molecular-weight 13157 #checksum 3535

Query Match 100.0%; Score 56; DB 13; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.17e+01;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 27 aaagilgtv 35  
|||||||  
QY 1 AAGIGILTV 9  
RESULT 2  
ENTRY I38506 #type complete  
TITLE melan-A protein - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change  
06-Sep-1996  
ACCESSIONS I38506  
REFERENCE I38506  
#authors Coullie, P.G.; Brichard, V.; Van pel, A.; Wolfel, T.;  
Schneider, J.; Traversari, C.; Mattei, S.; De Plaen, E.;  
Lurquin, C.; Szikora, J.P.; Renaud, J.; Boon, T.  
#journal J. Exp. Med. (1994) 180:35-42  
#title A new gene coding for a differentiation antigen recognized by  
autologous cytolytic T lymphocytes on HLA-A2 melanomas [see  
comments].

```

#cross-references MUID:94275389
#accession I38506
#status preliminary; translated from GB/EMBL/DBDJ
#molecule_type mRNA
#residues 1-118 #label RES
#cross-references EMBL:006654; NID:9517022; CDS_PID:9517023
SUMMARY #length 118 #molecular-weight 13157 #checksum 3535

Query Match 100.0%; Score 56; DB 13; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.17e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 aagigiltv 35
|||||
QY 1 AAGIGILTV 9

RESULT 3
ENTRY #type complete
TITLE H+-transporting ATP synthase (EC 3.6.1.34) proteolipid chain
ORGANISM - Sulfolobus acidocaldarius
DATE #formal_name Sulfolobus acidocaldarius
20-Dec-1989 #sequence_revision 20-Dec-1989 #text_change
23-Jun-1993
ACCESSIONS A33351
REFERENCE A33351
#authors Denda, K.; Konishi, J.; Oshima, T.; Date, T.; Yoshida, M.
#journal J. Biol. Chem. (1989) 264:7119-7121
#title A gene encoding the proteolipid subunit of Sulfolobus
acidocaldarius ATPase complex.
#cross-references MUID:8921412
#accession A33351
#status preliminary
#molecule_type DNA
#residues 1-101 #label DEN
#cross-references GB:J04740
KEYWORDS hydrolase
SUMMARY #length 101 #molecular-weight 10362 #checksum 4300

Query Match 83.9%; Score 47; DB 10; Length 101;
Best Local Similarity 87.5%; Pred. No. 8.60e+00;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 59 aagigvlt 66
|||||
QY 1 AAGIGILT 8

RESULT 4
ENTRY VGBEPB #type complete
TITLE glycoprotein gIII precursor - suid herpesvirus 1
ORGANISM #formal_name suid herpesvirus 1
DATE 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change
05-Jan-1996
ACCESSIONS A26097
REFERENCE A26097
#authors Robbins, A.K.; Watson, R.J.; Whealy, M.E.; Hays, W.W.;
Enquist, L.W.
#journal J. Virol. (1986) 58:339-347
#title Characterization of a Pseudorabies virus glycoprotein gene
with homology to herpes simplex virus type 1 and type 2
glycoprotein C.
#cross-references MUID:86200375
#accession A26097
#molecule_type DNA
#residues 1-479 #label ROB
#experimental_source strain Becker
CLASSIFICATION #superfamily herpesvirus glycoprotein F
KEYWORDS glycoprotein
FEATURE
1-22 #domain signal sequence #status predicted #label SIG
23-479 #product glycoprotein gIII #status predicted #label GPG
40,84,169,192,220,

228,285,302 #binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 479 #molecular-weight 51206 #checksum 1630

Query Match 83.9%; Score 47; DB 4; Length 479;
Best Local Similarity 75.0%; Pred. No. 8.60e+00;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigilai 463
|||||
QY 2 AGIGILTV 9

RESULT 5
ENTRY S48276 #type complete
TITLE YSA1 protein - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein YBR0907; protein YBR111c
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995
ACCESSIONS S48276; S45979; S25364; S44691
REFERENCE S48255
#authors Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann,
H.
#journal Yeast (1994) 10:1363-1381
#title Analysis of a 70 kb region on the right arm of yeast
chromosome II.
#accession S48276
#molecule_type DNA
#residues 1-231 #label MAN
#cross-references EMBL:X78993
#note nucleotide sequence is not given
REFERENCE S45927
#authors Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
#submission submitted to the Protein Sequence Database, August 1994
#accession S45979
#molecule_type DNA
#residues 1-231 #label FE2
#cross-references EMBL:X35980
REFERENCE S25364
#authors Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann,
H.
#journal Yeast (1992) 8:397-408
#title Molecular analysis of yeast chromosome II between CMD1 and
LYS2: the excision repair gene RAD16 located in this region
belongs to a novel group of double-finger proteins.
#accession S25364
#molecule_type DNA
#residues 1-47 #label MAW
#cross-references EMBL:X66247
GENETICS LISTA:YSA1
#gene
#map_position 2R
SUMMARY #length 231 #molecular-weight 26087 #checksum 4809

Query Match 82.1%; Score 46; DB 12; Length 231;
Best Local Similarity 85.7%; Pred. No. 1.35e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 79 gigiltv 85
|||||
QY 3 GIGILTV 9

RESULT 6
ENTRY S62369 #type complete
TITLE methylcobalamin: Coenzyme M methyltransferase (isoenzyme II)
- Methanosarcina barkeri
ORGANISM #formal_name Methanosarcina barkeri
DATE 20-Jul-1996 #sequence_revision 20-Jul-1996 #text_change
20-Jul-1996
ACCESSIONS S62369
REFERENCE S62368

```

```

#authors      Harms, U.; Thauer, R.K.
#journal      Eur. J. Biochem. (1996) 235:653-659
#title       Methylobacterales: coenzyme M methyltransferase isoenzymes MtaA
              and MtaB from Methanohalobium barkeri. Cloning, sequencing
              and differential transcription of the encoding genes, and
              functional overexpression of the mtaA gene in Escherichia
              coli.
#accession    S62369
#status       preliminary
#residues     1-339 ##label HAR
#cross-references EMBL:X91894
SUMMARY       #length 339 #molecular-weight 36761 #checksum 6431

Query Match   80.4%; Score 45; DB 16; Length 339;
Best Local Similarity 75.0%; Pred. No. 2.10e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 307 agvglltv 314
QY 2 AGIGILTV 9

RESULT 7
ENTRY   S59131 #type complete
TITLE   Kan-1 protein - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE     15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
         01-Mar-1996
ACCESSIONS S59131
REFERENCE   S59131
#authors   Furutani, M.; Arai, S.; Higashitsugu, H.; Wise, M.; Fukumoto,
              M.; Takano, S.; Nakayama, H.; Imamura, M.; Fujita, J.
#journal   Biochem. J. (1995) 311:203-208
#title     Reduced expression of kan-1 (encoding putative bile
              acid-CoA:amino acid N-acyltransferase) mRNA in livers of
              rats after partial hepatectomy and during sepsis.
#accession S59131
#status     preliminary
#molecule_type mRNA
#residues   1-420 ##label FUR
#cross-references EMBL:D43964
SUMMARY     #length 420 #molecular-weight 46496 #checksum 4868

Query Match   80.4%; Score 45; DB 14; Length 420;
Best Local Similarity 55.6%; Pred. No. 2.10e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 226 gpgvgilsv 234
QY 1 AAGIGILTV 9

RESULT 8
ENTRY   S19011 #type complete
TITLE   endo-1,4-beta-xylanase (EC 3.2.1.8) - Bacillus polymyxa
ORGANISM #formal_name Bacillus polymyxa
DATE     22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
         10-Nov-1995
ACCESSIONS S19011
REFERENCE   S19011
#authors   Gosalbes, M.J.; Perez-Gonzalez, J.A.; Gonzalez, R.; Navarro,
              A.
#journal   J. Bacteriol. (1991) 173:7705-7710
#title     Two beta-glycanase genes are clustered in Bacillus polymyxa:
              molecular cloning, expression, and sequence analysis of
              genes encoding a xylanase and an endo-beta-(1,3)-(1,
              4)-glucanase.
#cross-references MUID:92041687
#accession S19011
#status     preliminary
#molecule_type DNA
#residues   1-635 ##label GOS
#cross-references EMBL:X57094

```

```

#note         the authors translated the codon GAA for residue 78 as
              Gly, CCG for residue 272 as Thr, ATC for residue 412
              as Gln, and ATC for residue 478 as Tyr
KEYWORDS      glycosidase; hydrolase
SUMMARY       #length 635 #molecular-weight 67914 #checksum 2077

Query Match   78.6%; Score 44; DB 10; Length 635;
Best Local Similarity 75.0%; Pred. No. 3.24e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 149 gagigvlt 156
QY 1 AAGIGILTV 8

RESULT 9
ENTRY   S01903 #type complete
TITLE   H+-transporting ATP synthase (EC 3.6.1.34) epsilon chain -
              Arabidopsis thaliana chloroplast
ORGANISM #formal_name chloroplast Arabidopsis thaliana #common_name
              mouse-ear cress
DATE     01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change
         14-Sep-1994
ACCESSIONS S01903
REFERENCE   S01903
#authors   Chen, H.C.; Wintz, H.; Weill, J.H.; Pillay, D.T.N.
#journal   Nucleic Acids Res. (1988) 16:10372
#title     Nucleotide sequence of chloroplast CFI-ATPase epsilon-subunit
              and elongator tRNA(Met) genes from Arabidopsis thaliana.
#cross-references MUID:89057486
#accession S01903
#molecule_type DNA
#residues   1-132 ##label CHE
#cross-references EMBL:X12889
GENETICS      atpE
              #gene
              #genome
              #superfamily H+-transporting ATP synthase epsilon chain
              #ATP synthase; chloroplast; hydrolase; membrane-associated
              complex; thylakoid
KEYWORDS      #length 132 #molecular-weight 14472 #checksum 1607

Query Match   76.8%; Score 43; DB 5; Length 132;
Best Local Similarity 66.7%; Pred. No. 4.97e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 43 avdigilvti 51
QY 1 AAGIGILTV 9

RESULT 10
ENTRY   ACMSE #type complete
TITLE   nicotinic acetylcholine receptor epsilon chain precursor -
              mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
         27-Oct-1995
ACCESSIONS S13592; B33358
REFERENCE   S13592
#authors   Gardner, P.D.
#journal   Nucleic Acids Res. (1990) 18:6714
#title     Nucleotide sequence of the epsilon-subunit of the mouse
              muscle nicotinic acetylcholine receptor.
#cross-references MUID:91067487
#accession S13592
#molecule_type mRNA
#residues   1-493 ##label GAR
#cross-references EMBL:X55718
              A33358
#authors   Buonanno, A.; Mudd, J.; Merlie, J.P.
#journal   J. Biol. Chem. (1989) 264:7611-7616
#title     Isolation and characterization of the beta-and

```

```

epsilon-subunit genes of mouse muscle acetylcholine
receptor.
#cross-references MUID:89214211
#accession B3358
#molecule_type DNA
#residues 1-493 ##label BUO
#cross-references 1-493 ##label BUO
#cross-references GB:J04598
CLASSIFICATION #superfamily acetylcholine receptor
KEYWORDS glycoprotein; ion channel; muscle; neurotransmitter receptor;
postsynaptic membrane; transmembrane protein
FEATURE
1-20 #domain signal sequence #status predicted #label SIG\
21-493 #product nicotinic acetylcholine receptor epsilon chain
#status predicted #label MAT\
21-239 #domain extracellular #status predicted #label EXT\
240-266 #domain transmembrane #status predicted #label TM1\
273-291 #domain transmembrane #status predicted #label TM2\
307-328 #domain transmembrane #status predicted #label TM3\
329-456 #domain intracellular #status predicted #label INT\
457-479 #domain transmembrane #status predicted #label TM4\
86,161,327 #binding site carbohydrate (Asn) (covalent) #status
predicted\
148-162 #disulfide bonds #status predicted
SUMMARY #length 493 #molecular-weight 54914 #checksum 1794
Query Match 76.8%; Score 43; DB 3; Length 493;
Best Local Similarity 75.0%; Pred. No. 4.97e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 93 agvgilrv 100
||:||||
Qy 2 AAGIGILTV 9
RESULT 11
ENTRY #type complete
TITLE Hypothetical protein 68 - Herpetosiphon aurantiacus
ORGANISM #formal_name Herpetosiphon aurantiacus
DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
03-May-1994
ACCESSIONS JT0592; S21950; S21953
REFERENCE #authors Duesterhoeft, A.; Kroeger, M.
#journal Gene (1991) 106:87-92
#title Cloning, sequence and characterization of
mSc-methyltransferase-encoding gene, hgidiim (GTCGAC), from
Herpetosiphon giganteus strain Hpa2.
#cross-references MUID:92039068
#accession JT0592
#molecule_type DNA
#residues 1-611 ##label DUE
#cross-references EMBL:X55141
#experimental_source strain Hpa2
SUMMARY #length 611 #molecular-weight 68353 #checksum 9198
Query Match 76.8%; Score 43; DB 8; Length 611;
Best Local Similarity 71.4%; Pred. No. 4.97e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 97 gigilai 103
||:||||
Qy 3 GIGILTV 9
RESULT 12
ENTRY #type fragment
TITLE 31K flagellin: Methanococcus voltae (fragment)
ORGANISM #formal_name Methanococcus voltae
DATE 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change
23-Mar-1993
ACCESSIONS A34624
REFERENCE A34624
#authors Kalkmoff, M.L.; Karnauchow, T.M.; Jarrell, K.F.

```

```

#journal Biochem. Biophys. Res. Commun. (1990) 167:154-160
#title Conserved N-terminal sequences in the flagellins of
archaeobacteria.
#cross-references MUID:90179742
#accession A34624
#status preliminary
#molecule_type protein
#residues 1-19 ##label KAL
SUMMARY #length 19 #checksum 4377
Query Match 75.0%; Score 42; DB 10; Length 19;
Best Local Similarity 66.7%; Pred. No. 7.57e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 1 asgiqtliv 9
||:||||
Qy 1 AAGIGILTV 9
RESULT 13
ENTRY #type complete
TITLE VHBPD major capsid protein D - phage lambda
ALTERNATE_NAMES head protein D
ORGANISM #formal_name phage lambda
DATE 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change
17-Nov-1995
ACCESSIONS G04333; C43013; A04334; A23206
REFERENCE #authors Daniels, D.
#journal submitted to the Nucleic Acid Sequence Database, September
1982
#accession G04333
#molecule_type DNA
#residues 1-110 ##label DAN
REFERENCE #authors Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen,
G.B.
#journal J. Mol. Biol. (1982) 162:729-773
#title Nucleotide sequence of bacteriophage lambda DNA.
#cross-references MUID:83189071
#accession C43013
#molecule_type DNA
#residues 1-110 ##label SAN
REFERENCE #accession A23206
#authors Witkiewicz, H.; Schweiger, M.
#journal EMBO J. (1982) 1:1559-1564
#cross-references MUID:84207913
#contents annotation; physicochemical properties
COMMENT Gene D protein is a major component of the phage head and serves to
stabilize the head during DNA packaging. There are approximately
420 copies of protein D per mature phage.
GENETICS
#gene D
#map_position 11.85-12.53
CLASSIFICATION #superfamily phage lambda major capsid protein D
KEYWORDS DNA packaging
SUMMARY #length 110 #molecular-weight 11572 #checksum 3863
Query Match 75.0%; Score 42; DB 4; Length 110;
Best Local Similarity 55.6%; Pred. No. 7.57e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 56 gaavgilav 64
||:||||
Qy 1 AAGIGILTV 9
RESULT 14
ENTRY #type complete
TITLE flagellin B2 precursor - Methanococcus voltae
ORGANISM #formal_name Methanococcus voltae
DATE 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change
30-Sep-1993

```

```

ACCESSIONS      C41316
REFERENCE        A41316
#authors        Kalmokoff, M.L.; Jarrell, K.F.
#journal        J. Bacteriol. (1991) 173:7113-7125
#title          Cloning and sequencing of a multigene family encoding the
                flagellins of Methanococcus voltae.
#cross-references MUID:92041608
#accession      C41316
##status        preliminary
##molecule_type DNA
##residues      1-216 #label KAL
##cross-references GB:M72148
SUMMARY          #length 216 #molecular-weight 22799 #checksum 9101

Query Match      75.0%; Score 42; DB 10; Length 216;
Best Local Similarity 66.7%; Pred. No. 7.57e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 13 asgigtliiv 21
|:|||||
QY 1 AAGIGILTV 9

RESULT 15
ENTRY      B41316 #type complete
TITLE      flagellin B1 precursor - Methanococcus voltae
ORGANISM   #formal_name Methanococcus voltae
DATE       21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change
          30-Sep-1993
ACCESSIONS B41316
REFERENCE   A41316
#authors   Kalmokoff, M.L.; Jarrell, K.F.
#journal   J. Bacteriol. (1991) 173:7113-7125
#title     Cloning and sequencing of a multigene family encoding the
          flagellins of Methanococcus voltae.
#cross-references MUID:92041608
#accession B41316
##status   preliminary
##molecule_type DNA
##residues 1-218 #label KAL
##cross-references GB:M72148
SUMMARY     #length 218 #molecular-weight 22513 #checksum 7076

Query Match      75.0%; Score 42; DB 10; Length 218;
Best Local Similarity 66.7%; Pred. No. 7.57e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 13 asgigtliiv 21
|:|||||
QY 1 AAGIGILTV 9

```

Search completed: Tue Jun 10 11:17:45 1997  
 Job time : 11 secs.

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Blocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:17:05 1997; MasPar time 2.05 Seconds  
Tabular output not generated. 93.292 Million cell updates/sec

Title: >US-08-231-565A-4  
Description: (1-9) from US08231565A.pep  
Perfect Score: 56  
Sequence: 1 AAGIGILT 9

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 22.471; Variance 24.207; scale 0.928

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	47	83.9	101	1	ATPL_SULAC MEMBRANE-ASSOCIATED A	2.73e+00
2	47	83.9	479	10	VGLC_PRVIF GLYCOPROTEIN GIII PRE	2.73e+00
3	46	82.1	231	11	YSAL_YEAST YSAL PROTEIN.	4.59e+00
4	44	78.6	456	4	GLMU_ECOLI UDP-N-ACETYLGLUCOSAMI	1.26e+01
5	44	78.6	635	10	XVND_BACPO ENDO-1,4-BETA-XYLANS	1.26e+01
6	43	76.8	132	1	ATPE_ARATH ATP SYNTHASE EPSILON	2.05e+01
7	43	76.8	388	11	YXBC_BACSU HYPOTHETICAL METABOLI	2.05e+01
8	43	76.8	493	1	ACHE_MOUSE ACETYLCHOLINE RECEPTO	2.05e+01
9	43	76.8	611	11	YD3M_HERAU HYPOTHETICAL 68.4 KD	2.05e+01
10	42	75.0	110	10	VCAD_LAMB2 HEAD DECORATION PROTE	3.32e+01
11	42	75.0	216	4	FLA2_METVO FLAGELLIN B2 PRECURSO	3.32e+01
12	42	75.0	218	4	FLA1_METVO FLAGELLIN B1 PRECURSO	3.32e+01
13	42	75.0	308	6	MENA_HAEIN 1,4-DIHYDROXY-2-NAPHT	3.32e+01
14	42	75.0	332	1	ACOA_ALCEU ACETOIN:2,6-DICHLOROP	3.32e+01
15	42	75.0	461	9	THDF_HAEIN POSSIBLE THIOPHENE AN	3.32e+01
16	42	75.0	501	6	LYSL_CORGL L-LYSINE TRANSPORT PR	3.32e+01
17	42	75.0	530	1	AIP2_YEAST ACTIN INTERACTING PRO	3.32e+01
18	42	75.0	659	11	YIFB_BACSU HYPOTHETICAL 74.3 KD	3.32e+01
19	42	75.0	1530	1	BFRI_SCHPO BREFFELDIN A RESISTANC	3.32e+01
20	41	73.2	225	2	CD9_FELCA CD9 ANTIGEN.	5.32e+01
21	41	73.2	225	2	CD9_BOVIN CD9 ANTIGEN.	5.32e+01
22	41	73.2	271	11	YK23_YEAST HYPOTHETICAL 31.0 KD	5.32e+01

23	41	73.2	345	5	IAP_ECOLI ALKALINE PHOSPHATASE	5.32e+01
24	41	73.2	348	9	SELD_HAEIN SELENIDE/WATER DIKINA	5.32e+01
25	41	73.2	359	11	YFDA_CORGL HYPOTHETICAL PROTEIN	5.32e+01
26	41	73.2	401	10	YABA_SCHPO HYPOTHETICAL 44.4 KD	5.32e+01
27	41	73.2	666	2	COL4_BRAJA PROBABLE CYTOCHROME C	5.32e+01
28	41	73.2	752	6	METE_ECOLI 5-METHYLATETRAHYDROPE	5.32e+01
29	41	73.2	845	6	MAT3_RAT MATRIN 3.	5.32e+01
30	41	73.2	977	11	YDGB_SCHPO HYPOTHETICAL 111.4 KD	5.32e+01
31	40	71.4	75	4	GVPI2_HALHA GAS VESICLE PROTEIN,	8.44e+01
32	40	71.4	78	4	GVPI2_HALHA GAS VESICLE PROTEIN,	8.44e+01
33	40	71.4	78	4	GVPA_HALME GAS VESICLE PROTEIN (	8.44e+01
34	40	71.4	219	4	FLAA_METVO FLAGELLIN A PRECURSOR	8.44e+01
35	40	71.4	231	1	ATMC_SALTY MG(2+) TRANSPORT ATPA	8.44e+01
36	40	71.4	251	10	VP8_ECOLI HYPOTHETICAL 35.0 KD	8.44e+01
37	40	71.4	321	11	YHBE_ECOLI STRUCTURAL PROTEIN VP	8.44e+01
38	40	71.4	461	6	NIFN_KLEPN NITROGENASE IRON-MOLY	8.44e+01
39	40	71.4	491	1	AMPG_ECOLI AMPG PROTEIN.	8.44e+01
40	40	71.4	566	6	NARQ_ECOLI NITRATE/NITRITE SENSO	8.44e+01
41	40	71.4	590	2	CHLI_ARATH NITRATE/CHLORATE TRAN	8.44e+01
42	40	71.4	591	11	YOO1_BPL2 HYPOTHETICAL 66.6 KD	8.44e+01
43	40	71.4	626	6	MAOM_SOLTU MALATE OXIDOREDUCTASE	8.44e+01
44	40	71.4	2051	3	FAS1_YEAST FATTY ACID SYNTHASE.	8.44e+01
45	39	69.6	236	1	BACT_HALVA SENSORY RHODOPSIN II	1.32e+02

ALIGNMENTS

RESULT 1  
ID ATPL\_SULAC STANDARD; PRT; 101 AA.  
AC P23040;  
DT 01-NOV-1991 (REL. 20, CREATED)  
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE MEMBRANE-ASSOCIATED ATPASE C CHAIN (EC 3.6.1.34) (SUL-ATPASE  
DE PROTEOLIPID CHAIN).  
GN ATPP.  
OS SULFOLOBUS ACIDOCALDARIUS.  
OC ARCHAEABACTERIA; CRENARCHAEOTA; SULFOLOBALES.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 89214142.  
RA DENDA K., KONISHI J., OSHIMA T., DATE T., YOSHIDA M.;  
RL J. BIOL. CHEM. 264:7119-7121(1989).  
CC !- FUNCTION: THE C CHAIN IS A PROPEOLIPID, AND ONE OF THE MEMBRANOUS  
CC SUBUNITS OF THE THE NONENZYMATIC COMPONENT OF THE SUL-ATPASE  
CC COMPLEX.  
CC !- SUBUNIT: SUL-ATPASE IS COMPOSED OF SIX (OR FIVE ?) SUBUNITS:  
CC ALPHA, BETA, DELTA, GAMMA, C (PROTEOLIPID), AND POSSIBLY EPSILON.  
CC !- SIMILARITY: TO THE PROTEOLIPID SUBUNIT OF FOF1-ATPASES.  
DR EMBL; J04740; G152925; -  
DR PIR; A33351; A33351.  
KW HYDROGEN ION TRANSPORT; LIPID-BINDING; TRANSMEMBRANE.  
FT TRANSMEM 5 25 POTENTIAL.  
FT TRANSMEM 37 57 POTENTIAL.  
FT TRANSMEM 75 95 POTENTIAL.  
SQ SEQUENCE 101 AA; 10362 MW; 1DC8C74D CRC32;

Query Match 83.9%; Score 47; DB 1; Length 101;  
Best Local Similarity 87.5%; Pred. No. 2.73e+00;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 59 aaagivlt 66  
QY 1 AAGIGILT 8  
|||||:||

RESULT 2  
ID VGLC\_PRVIF STANDARD; PRT; 479 AA.  
AC P06024;  
DT 13-AUG-1987 (REL. 05, CREATED)  
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
DE GLYCOPROTEIN GIII PRECURSOR.



OS PSEUDORABIES VIRUS (STRAIN INDIANA-FUNKHAUSER / BECKER) (PRV).  
 CC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAPERESVIRINAE.

CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN C.  
 DR EMBL; M12778; G334050; -.  
 KW PIR; A26097; VGBEPB.  
 DR GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.  
 FT SIGNAL 1 22  
 FT CHAIN 23 479 GLYCOPROTEIN GIII.  
 FT CARBOHYD 40 40 POTENTIAL.  
 FT CARBOHYD 40 40 POTENTIAL.  
 FT CARBOHYD 84 84 POTENTIAL.  
 FT CARBOHYD 169 169 POTENTIAL.  
 FT CARBOHYD 192 192 POTENTIAL.  
 FT CARBOHYD 220 220 POTENTIAL.  
 FT CARBOHYD 228 228 POTENTIAL.  
 FT CARBOHYD 285 285 POTENTIAL.  
 FT CARBOHYD 302 302 POTENTIAL.  
 SQ SEQUENCE 479 AA; 51206 MW; 42EE5703 CRC32;

Query Match 83.9%; Score 47; DB 10; Length 479;  
 Best Local Similarity 75.0%; Pred. No. 2.73e+00;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigilal 463  
 |||||  
 QY 2 AGIGILTV 9

RESULT 3  
 ID YSAI\_YEAST STANDARD; PRT; 231 AA.  
 AC Q01976;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE YSAI PROTEIN.  
 GN YSAI OR YBR111C OR YBR0907.  
 OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).  
 CC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE; 95208357.  
 RA MANNAHAUT G., STUCKA R., EHNLE S., VETTER I., FELDMANN H.;  
 RL YEAST 10:1363-1381(1994).  
 RN [2]  
 RP SEQUENCE OF 1-47 FROM N.A.  
 RC STRAIN-S288C;  
 RX MEDLINE; 92327848.  
 RA MANNAHAUT G., STUCKA R., EHNLE S., VETTER I., FELDMANN H.;  
 RL YEAST 8:397-408(1992).  
 CC -1- SIMILARITY: STRONG, TO B. SUBTILIS YOKG.  
 DR EMBL; Z35980; G536466; -.  
 DR EMBL; X78993; G476067; -.  
 DR EMBL; X66247; G3549; -.  
 DR PIR; S44691; S44691.  
 DR LISTA; SC01415; YSAI.  
 DR SGD; L0002551; YSAI.  
 DR PROSITE; PS00893; MUTT.  
 FT DOMAIN 112 145 MUTT-LIKE.  
 SQ SEQUENCE 231 AA; 26087 MW; 49A2D6CB CRC32;

Query Match 82.1%; Score 46; DB 11; Length 231;  
 Best Local Similarity 85.7%; Pred. No. 4.59e+00;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 79 gigilitv 85  
 |||||  
 QY 2 AGIGILTV 9

QY 3 CIGILTV 9  
 RESULT 4  
 ID GLMU\_ECOLI STANDARD; PRT; 456 AA.  
 AC P17114;  
 DT 01-AUG-1990 (REL. 15, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE UDP-N-ACETYLGLUCOSAMINE PYROPHOSPHORYLASE (EC 2.7.7.23) (N-ACETYLGLUCOSAMINE-1-PHOSPHATE URIDYLTRANSFERASE).  
 GN GLMU.  
 OS ESCHERICHIA COLI.  
 CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS; ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 85121806.  
 RA WALKER J.E., GAY N.J., SARASTE M., EBERLE A.N.;  
 RL BIOCHEM. J. 224:799-815(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE; 93315143.  
 RA BURLAND V.D., PLUNKETT G. III, DANIELS D.L., BLATTNER F.R.;  
 RL GENOMICS 16:551-561(1993).  
 RN [3]  
 RP IDENTIFICATION.

RX MENGIN-LECEULX D., VAN HEIJENCOORT J.;  
 RL J. BACTERIOL. 175:6150-6157(1993).  
 CC -1- FUNCTION: BIFUNCTIONAL ENZYME RESPONSIBLE FOR THE ACETYLATION OF GLC-N-1-P TO GIVE GLCNAC-1-P AND THE SYNTHESIS OF UDP-GLCNAC.  
 CC -1- CATALYTIC ACTIVITY: UTP + N-ACETYL-ALPHA-D-GLUCOSAMINE 1-PHOSPHATE -> PYROPHOSPHATE + UDP-N-ACETYL-D-GLUCOSAMINE.  
 CC -1- PATHWAY: PEPTIDOGLYCAN AND LIPOPOLYSACCHARIDE BIOSYNTHESIS.  
 CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT THAT CREATES TWO ORFS.  
 CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).  
 DR EMBL; X01631; G43267; -.  
 DR EMBL; L10328; G290579; ALT\_FRAME.  
 DR EMBL; L10328; G290578; ALT\_FRAME.  
 DR EMBL; G811198; GLMU.  
 DR PROSITE; PS00101; HEXAPEPT\_TRANSFERASES.  
 KW PEPTIDOGLYCAN SYNTHESIS; CELL WALL; TRANSFERASE;  
 KW NUCLEOTIDYLTRANSFERASE; REPEAT; MULTIFUNCTIONAL ENZYME.  
 SQ SEQUENCE 456 AA; 49162 MW; 7A80D509 CRC32;

Query Match 78.6%; Score 44; DB 4; Length 456;  
 Best Local Similarity 75.0%; Pred. No. 1.26e+01;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 124 ggigilv 131  
 |||||  
 QY 2 AGIGILTV 9

RESULT 5  
 ID XYND\_BACPO STANDARD; PRT; 635 AA.  
 AC P45796;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE ENDO-1,4-BETA-XYLANASE D PRECURSOR (EC 3.2.1.8) (XYLANASE D)  
 DE (1,4-BETA-D-XLAN XYLANOHYDROLASE D).  
 GN XYND.

OS BACILLUS POLYMYXA.  
 CC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 842;  
 RX MEDLINE; 92041687.

RA GOSALBES M.J., PEREZ-GONZALEZ J.A., GONZALEZ R., NAVARRO A.;  
 RL J. BACTERIOL. 173:7705-7710(1991).  
 CC -!- FUNCTION: SHOWS XYLANASE ACTIVITY AS WELL AS ALPHA-L-  
 CC ARABINOFURANOSIDASE ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: ENDHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC  
 CC LINKAGES IN XYLANS.  
 CC -!- PATHWAY: XYLAN DEGRADATION.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 43 OF GLYCOSYL HYDROLASES.  
 DR EMBL; X57094; G48816; -.  
 CC XYLAN DEGRADATION; HYDROLASE; GLYCOSIDASE; SIGNAL.  
 FT SIGNAL 1 26  
 FT CHAIN 27 635 XYLANASE D.  
 SQ SEQUENCE 635 AA; 67914 MW; 078AAB82 CRC32;  
 Query Match 78.6%; Score 44; DB 10; Length 635;  
 Best Local Similarity 75.0%; Pred. No. 1.26e+01;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Db 149 gagigvlt 156  
 QY :|||||  
 1 AAGIGILT 8  
 RESULT 6  
 ID ATPE\_ARATH STANDARD; PRT; 132 AA.  
 AC P09468;  
 DT 01-MAR-1989 (REL. 10, CREATED)  
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE ATP SYNTHASE EPSILON CHAIN (EC 3.6.1.34).  
 GN ATPE.  
 OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
 OG CHLOROPLAST.  
 CC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
 CC CAPPARALES; CRUCIFERAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. LANDSBERG ERECTA;  
 RX MEDLINE; 89057486.  
 RA CHEN H.-C., WINTZ H., WEIL J.-H., PILLAY D.T.N.;  
 RL NUCLEIC ACIDS RES. 16:10372-10372(1988).  
 CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
 CC GRADIENT ACROSS THE MEMBRANE.  
 CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
 CC HAS THREE MAIN SUBUNITS: A, B AND C.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.  
 CC -!- SIMILARITY: BELONGS TO THE ATPASE EPSILON CHAIN FAMILY.  
 DR EMBL; X12889; G11334; -.  
 DR PIR; S01903; S01903.  
 KW ATP SYNTHESIS; CHLOROPLAST; THYLAKOID MEMBRANE; CF(1);  
 SQ HYDROLASE; HYDROGEN ION TRANSPORT.  
 Query Match 76.8%; Score 43; DB 1; Length 132;  
 Best Local Similarity 66.7%; Pred. No. 2.05e+01;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Db 43 avdigilti 51  
 QY :|||||  
 1 AAGIGILT 9  
 RESULT 7  
 ID YXBC\_BACSU STANDARD; PRT; 388 AA.  
 AC P46333;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN IN HTPG-IOLR INTERGENIC  
 DE REGION.  
 GN YXBC OR SS92BR.

OS BACILLUS SUBTILIS.  
 CC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / BGSC1A1;  
 RX MEDLINE; 96093926.  
 RA YOSHIDA K.-I., SEKI S., FUJIMURA M., MIWA Y., YOSHIDA K.-I.;  
 RL DNA RES. 2:61-69(1995).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
 DR EMBL; D45242; G904202; -.  
 DR SUBTILIST; BG11360; YXBC.  
 DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1.  
 DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2.  
 KW HYPOTHETICAL PROTEIN; TRANSPORT; TRANSMEMBRANE.  
 FT TRANSMEM 15 35 POTENTIAL.  
 FT TRANSMEM 39 59 POTENTIAL.  
 FT TRANSMEM 77 97 POTENTIAL.  
 FT TRANSMEM 105 125 POTENTIAL.  
 FT TRANSMEM 140 160 POTENTIAL.  
 FT TRANSMEM 164 184 POTENTIAL.  
 FT TRANSMEM 242 262 POTENTIAL.  
 FT TRANSMEM 281 301 POTENTIAL.  
 FT TRANSMEM 309 329 POTENTIAL.  
 FT TRANSMEM 342 362 POTENTIAL.  
 SQ SEQUENCE 388 AA; 42163 MW; 7E9AEDF4 CRC32;  
 Query Match 76.8%; Score 43; DB 11; Length 388;  
 Best Local Similarity 85.7%; Pred. No. 2.05e+01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 285 gigilnv 291  
 QY :|||||  
 3 GIGILT 9  
 RESULT 8  
 ID ACHE\_MOUSE STANDARD; PRT; 493 AA.  
 AC P20782;  
 DT 01-FEB-1991 (REL. 17, CREATED)  
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE ACETYLCHOLINE RECEPTOR PROTEIN, EPSILON CHAIN PRECURSOR.  
 GN ACHE.  
 OS MUS MUSCULUS (MOUSE).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 CC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 89214211.  
 RA BUONANNO A., MUDD J., MERLIE J.P.;  
 RL J. BIOL. CHEM. 264:7611-7616(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91067487.  
 RA GARDNER P.D.;  
 RL NUCLEIC ACIDS RES. 18:6714-6714(1990).  
 CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN  
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND  
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA  
 CC MEMBRANE.  
 CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,  
 CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE  
 CC MUSCLE) CHAINS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 DR EMBL; J04698; G387086; -.  
 DR EMBL; X55718; G53161; -.  
 DR PIR; S13592; ACMSE.  
 DR PIR; B33358; B33358.  
 DR PROSITE; PS00236; NEUROTR\_ION\_CHANNEL.  
 KW RECEPTOR; POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN; SIGNAL;  
 KW TRANSMEMBRANE.

```

FT SIGNAL      1 20
FT CHAIN      21 493
FT DOMAIN     21 239
FT TRANSMEM   240 264
FT TRANSMEM   265 272
FT TRANSMEM   273 291
FT TRANSMEM   292 306
FT TRANSMEM   307 328
FT TRANSMEM   329 456
FT TRANSMEM   457 480
FT TRANSMEM   481 493
FT DISULFID   148 162
FT CARBOHYD   86 86
FT CARBOHYD   161 161
SQ SEQUENCE   493 AA; 54914 MW; BB5BF2C0 CRC32;

Query Match      76.8%; Score 43; DB 1; Length 493;
Best Local Similarity 75.0%; Pred. No. 2.05e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 93 agvgilrv 100
QY 2 AGIGILTV 9

RESULT 9
ID YD3M.HERAU STANDARD; PRT; 611 AA.
AC P25280;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DE HYPOHETICAL 68.4 KD PROTEIN IN HGDIIM 3'REGION (ORF68).
OS HERPETOIPHON AURANTICUS (HERPETOIPHON GIGANTEUS).
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;
OC NONPHOTOSYNTHETIC, NONFRUITING GLIDING; BEGGIATOACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HPA2;
RX MEDLINE; 92039068.
RL GENE 106:87-92(1991).
DR EMBL; X55141; G48774; -.
DR PIR; J05092; J05092.
DR PIR; S21953; S21953.
DR PIR; S21950; S21950.
KW HYPOHETICAL PROTEIN; RESTRICTION SYSTEM; REPEAT.
FT DOMAIN     382 403
FT REPEAT     382 392
FT REPEAT     393 403
FT REPEAT     404 409
SQ SEQUENCE   611 AA; 68354 MW; 473CD6A4 CRC32;

Query Match      76.8%; Score 43; DB 11; Length 611;
Best Local Similarity 71.4%; Pred. No. 2.05e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 97 gigitlrv 103
QY 3 GIGILTV 9

RESULT 10
ID VCAD_LAMBDA STANDARD; PRT; 110 AA.
AC P03712;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE HEAD DECORATION PROTEIN (GPD) (MAJOR CAPSID PROTEIN D).
GN D.
OS BACTERIOPHAGE LAMBDA.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; SIPHOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE; 83189071.
RA SANGER F., COULSON A.R., HONG G.F., HILL D.F., PETERSEN G.B.;
J. MOL. BIOL. 162:729-773(1982).
RN [2]
RX SEQUENCE.
RX MEDLINE; 84207913.
RA WITKIEWICZ H., SCHWEIGER M.;
EMBO J. 1:1559-1564(1982).
CC -I- FUNCTION: STABILIZES THE HEAD SHELL FOLLOWING THE REARRANGEMENT
OF THE GPE SUBUNITS OF THE HEAD SHELL LATTICE THAT ACCOMPANIES
EXPANSION OF THE HEAD. THERE ARE APPROXIMATELY 420 COPIES OF
PROTEIN D PER MATURE PHAGE.
CC -I- SIMILARITY: TO BACTERIOPHAGE 21 HEAD DECORATION PROTEIN.
DR EMBL; J02459; G215111; -.
DR PIR; A04334; VHEPDL.
DR PIR; A23206; A23206.
KW COAT PROTEIN.
SQ SEQUENCE 110 AA; 11572 MW; FDD50011 CRC32;

Query Match      75.0%; Score 42; DB 10; Length 110;
Best Local Similarity 55.6%; Pred. No. 3.32e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 56 gaavgilav 64
QY 1 AAGIGILTV 9

RESULT 11
ID FLA2.METVO STANDARD; PRT; 216 AA.
AC P27804; P17602;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE FLAGELLIN B2 PRECURSOR.
GN FLA2.
OS METHANOCOCCUS VOLTAE.
OC ARCHAEABACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PS;
RX MEDLINE; 92041608.
RA KALMOKOFF M.L., JARRELL K.F.;
RL J. BACTERIOL. 173:7113-7125(1991).
RN [2]
RP SEQUENCE OF 13-32.
RC STRAIN=PS;
RX MEDLINE; 90179742.
RA KALMOKOFF M.L., KARNAUCHOW T.M., JARRELL K.F.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 167:154-160(1990).
CC -I- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -I- SIMILARITY: TO OTHER ARCHAEABACTERIAL FLAGELLINS.
DR EMBL; M72148; G150063; -.
DR PIR; A34624; A34624.
DR PIR; C41316; C41316.
KW FLAGELLA; MULTIGENE FAMILY.
FT PROPEP     1 12
FT CHAIN      13 216
SQ SEQUENCE   216 AA; 22799 MW; B62A0B23 CRC32;

Query Match      75.0%; Score 42; DB 4; Length 216;
Best Local Similarity 66.7%; Pred. No. 3.32e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 13 asgigtllv 21
QY 1 AAGIGILTV 9

RESULT 12
ID FLA1.METVO STANDARD; PRT; 218 AA.
AC P27803; P17602;

```

```

DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE FLAGELLIN B1 PRECURSOR.
GN FLAB1.
OS METHANOCOCCUS VOLTAE.
OC ARCHAEABACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PS;
RX MEDLINE; 92041608.
RA KALMOKOFF M.L., JARRELL K.F.;
RL J. BACTERIOL. 173:7113-7125(1991).
RN [2]
RP SEQUENCE OF 13-32.
RC STRAIN-PS;
RX MEDLINE; 90179742.
RA KALMOKOFF M.L., KARNAUCHOW T.M., JARRELL K.F.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 167:154-160(1990).
CC -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -!- SIMILARITY: TO OTHER ARCHAEABACTERIAL FLAGELLINS.
DR EMBL; M72148; G150062; -.
DR PIR; A34624; A34624.
DR PIR; B41316; B41316.
KW FLAGELLA; MULTIGENE FAMILY.
FT PROPEP 1
FT CHAIN 13 218 FLAGELLIN B1.
SQ SEQUENCE 218 AA; 22513 MW; E75F36F3 CRC32;

Query Match 75.0%; Score 42; DB 4; Length 218;
Best Local Similarity 66.7%; Pred. No. 3.32e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 13 asgigtliiv 21
QY 1 AAGIGILTV 9

RESULT 13
ID MENA_HAEIN STANDARD; PRT; 308 AA.
AC P44739;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE 1,4-DIHYDROXY-2-NAPHTHOATE OCTAPENYLTRANSFERASE (EC 2.5.1.11).
GN MENA OR HI0509.
OS HAEMOPHILUS INFLUENZAE.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC PASTURELLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA FLEISCHMANN R.D., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA KERLAVAGE A.R., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., KELLELY J.M.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., COTTON M.D.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RL SCIENCE 269:496-512(1995).
CC -!- FUNCTION: CONVERSION OF 1,4-DIHYDROXY-2-NAPHTHOATE TO DIMETHYL-
CC MENAQUINONE (BY SIMILARITY).
CC -!- PATHWAY: MENAQUINONE BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: TO B.SUBTILIS YWAB (IPA-6D).
DR EMBL; L45150; G1003902; -.
DR EMBL; U32732; G925532; -.
KW MENAQUINONE BIOSYNTHESIS; TRANSFERASE; TRANSMEMBRANE.
FT TRANSMEM 22 42
POTENTIAL.

```

```

FT TRANSMEM 47 67
FT TRANSMEM 101 121
FT TRANSMEM 129 149
FT TRANSMEM 153 173
FT TRANSMEM 186 206
FT TRANSMEM 235 255
FT TRANSMEM 286 306
SQ SEQUENCE 308 AA; 33345 MW; 090B2655 CRC32;

Query Match 75.0%; Score 42; DB 6; Length 308;
Best Local Similarity 62.5%; Pred. No. 3.32e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 132 aglgilal 139
QY 2 AGIGILTV 9

RESULT 14
ID ACOA_ALCEU STANDARD; PRT; 332 AA.
AC P27745;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE ACETOIN:2,6-DICHLOROPHENOLINDOPHENOL OXIDOREDUCTASE ALPHA
DE SUBUNIT (EC 1.1.1.1.-) (ACETOIN:DCPIP OXIDOREDUCTASE-ALPHA)
DE (AO:DCPIP OR).
GN ACOA.
OS ALCALIGENES EUTROPHUS.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC ALCALIGENACEAE.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-31.
RC STRAIN-H16;
RX MEDLINE; 91286190.
RA PRIEFERT H., HEIN S., KRUEGER N., ZEH K., SCHMIDT B.,
RA STEINBUCHER A.;
RL J. BACTERIOL. 173:4056-4071(1991).
CC -!- FUNCTION: CATALYZES THE 2,6-DICHLOROPHENOLINDOPHENOL-DEPENDENT
CC CLEAVAGE OF ACETOIN INTO ACETATE AND ACETALDEHYDE IN VITRO. THE
CC ALPHA SUBUNIT IS PROBABLY THE CATALYTIC SUBUNIT OF THE ENZYME.
CC -!- PATHWAY: ACETOIN CATABOLISM.
CC -!- COFACTOR: THIAMINE PYROPHOSPHATE.
CC -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -!- SIMILARITY: TO THE ALPHA SUBUNITS OF 2-OXO-ACID DEHYDROGENASE
CC COMPONENTS OF VARIOUS MULTIENZYME COMPLEXES.
DR EMBL; M66060; G141894; -.
DR PIR; B42462; DEALXE.
KW OXIDOREDUCTASE; FLAVOPROTEIN; THIAMINE PYROPHOSPHATE.
FT INIT_MET 0
FT BINDING 173 173 THIAMINE PYROPHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 332 AA; 35243 MW; 3322DA8E CRC32;

Query Match 75.0%; Score 42; DB 1; Length 332;
Best Local Similarity 85.7%; Pred. No. 3.32e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 56 aagvgil 62
QY 1 AAGIGIL 7

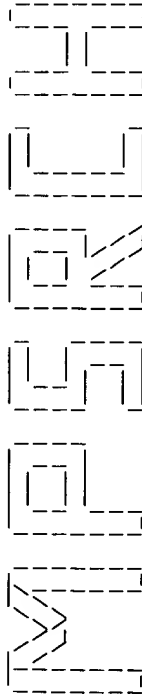
RESULT 15
ID THDF_HAEIN STANDARD; PRT; 461 AA.
AC P43730;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE POSSIBLE THIOPHENE AND FURAN OXIDATION PROTEIN THDF.
GN THDF OR HI1002.
OS HAEMOPHILUS INFLUENZAE.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC PASTURELLACEAE.

```

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RD / KW20;  
RX MEDLINE; 95350630.  
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,  
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,  
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,  
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,  
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
RA VENTER J.C.;  
RL SCIENCE 269:496-512(1995).  
CC -!- FUNCTION: INVOLVED IN THIOPHENE OXIDATION (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE ERA/THDF FAMILY OF GTP-BINDING  
CC PROTEINS.  
DR EMBL; L45640; G1006205; -.  
DR EMBL; U32781; G926061; -.  
KW GTP-BINDING.  
FT NP\_BIND 230 237 GTP (BY SIMILARITY).  
FT NP\_BIND 277 281 GTP (BY SIMILARITY).  
FT NP\_BIND 342 345 GTP (BY SIMILARITY).  
SQ SEQUENCE 461 AA; 50446 MW; 66E6E93F CRC32;  
  
Query Match 75.0%; Score 42; DB 9; Length 461;  
Best Local Similarity 75.0%; Pred.No. 3.32e+01;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Db 24 ggigilrv 31  
:|||||  
QY 2 AGIGILTV 9

Search completed: Tue Jun 10 11:17:16 1997  
Job time : 11 secs.

\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, Blocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:19:21 1997; MasPar time 1.99 Seconds  
Tabular output not generated. 54.816 Million cell updates/sec

Title: >US-08-231-565A-17  
Description: (1-10) from US08231565A.pep  
Perfect Score: 62  
Sequence: 1 EAAGIGILTV 10

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 15.466; Variance 45.800; scale 0.338

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	62	100.0	10 15	R84197	MART-1 melanoma antig	1.03e+00
2	62	100.0	118 12	R63158	Tumour rejection anti	1.03e+00
3	62	100.0	118 15	R84212	MART-1 melanoma anti	1.03e+00
4	56	90.3	9 15	R84196	MART-1 melanoma anti	5.22e+00
5	56	90.3	10 15	R84198	MART-1 melanoma anti	5.22e+00
6	52	83.9	9 15	R84788	Modified MART-1 melan	1.50e+01
7	52	83.9	9 15	R84787	Modified MART-1 melan	1.50e+01
8	52	83.9	9 15	R84786	Modified MART-1 melan	1.50e+01
9	51	82.3	9 15	R84785	Modified MART-1 melan	1.95e+01
10	50	80.6	9 15	R84784	Modified MART-1 melan	2.53e+01
11	49	79.0	9 15	R84783	Modified MART-1 melan	3.27e+01
12	48	77.4	9 15	R84792	Modified MART-1 melan	4.23e+01
13	48	77.4	9 15	R84764	MART-1 melanoma anti	4.23e+01
14	48	77.4	9 15	R84796	Modified MART-1 melan	4.23e+01
15	48	77.4	9 15	R84795	Modified MART-1 melan	4.23e+01
16	48	77.4	9 15	R84793	Modified MART-1 melan	4.23e+01
17	48	77.4	9 15	R84794	Modified MART-1 melan	4.23e+01
18	48	77.4	9 15	R84797	Modified MART-1 melan	4.23e+01
19	47	75.8	479 3	P50034	Sequence encoded by t	5.46e+01
20	47	75.8	479 1	P81013	Complete sequence of	5.46e+01

21	43	69.4	9 15	R84790	Modified MART-1 melan	1.49e+02
22	43	69.4	231 5	R28440	Cephalosporin C acety	1.49e+02
23	43	69.4	1080 3	P50296	Mouse epidermal growth	1.49e+02
24	42	67.7	10 15	R84775	MART-1 melanoma anti	1.90e+02
25	41	66.1	9 15	R84791	Modified MART-1 melan	2.43e+02
26	41	66.1	9 15	R84800	Modified MART-1 melan	2.43e+02
27	41	66.1	9 15	R84789	Modified MART-1 melan	2.43e+02
28	41	66.1	9 15	R84798	Modified MART-1 melan	2.43e+02
29	41	66.1	9 15	R84799	Modified MART-1 melan	2.43e+02
30	41	66.1	25 8	R39763	Melittin delta 4 (del	2.43e+02
31	41	66.1	288 12	R66467	ALL-1 breakpoint clus	2.43e+02
32	41	66.1	1464 18	R80970	Human excitatory amin	2.43e+02
33	41	66.1	1464 12	R66039	Human N-methyl-D-aspa	2.43e+02
34	41	66.1	1464 10	R55529	Human NMDA R2A recept	2.43e+02
35	41	66.1	3910 7	R38470	ALL-1 protein.	2.43e+02
36	41	66.1	3910 12	R66462	ALL-1 (acute lymphocy	2.43e+02
37	41	66.1	3969 10	R52971	Product of the cDNA e	2.43e+02
38	40	64.5	185 8	R47115	Toxoplasma GP28.5 ant	3.10e+02
39	40	64.5	410 13	R69555	Human lysosomal membr	3.10e+02
40	40	64.5	542 16	R90772	Bacillus stearothermo	3.10e+02
41	40	64.5	1087 17	R96028	P. gingivalis haemagg	3.10e+02
42	40	64.5	1358 17	R96032	P. gingivalis haemagg	3.10e+02
43	40	64.5	1687 17	R96033	P. gingivalis haemagg	3.10e+02
44	40	64.5	1704 13	R70188	Arg-gingipain-2 prepo	3.10e+02
45	40	64.5	1732 17	R96029	P. gingivalis porphyp	3.10e+02

ALIGNMENTS

RESULT 1  
ID R84197 standard; Peptide; 10 AA.  
AC R84197;  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen immunogenic peptide M10-3 derivative.  
KW MART-1; M10-3; melanoma antigen recognised by T-cells; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN WO93P9193-A2.  
PD 02-NDV-1995.  
PF 21-APR-1995; U05063.  
PR 32-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 12; Page 122; 184pp; English.  
CC Immunogenic peptide M10-3 is a derivative of peptide M9-2 (R84196)  
CC which is based on the melanoma antigen (MART-1) (see R84212).  
CC M9-2 may be modified to improve immunogenicity (see R84783-R84800)  
CC and used in medicaments for the treatment or prevention (by  
CC immunization) of melanoma. Antibodies against MART-1 and its  
CC immunogenic peptides may be used in the detection and isolation of  
CC MART-1 from a sample, the detection of which is indicative of a  
CC disease state (melanoma or metastatic melanoma).  
CC See also R84198  
SQ Sequence 10 AA;

Query Match 100.0%; Score 62; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.03e+00;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 eaagigiltv 10  
| | | | | | | | | |  
QY 1 EAAGIGILTV 10

RESULT 2  
ID R63158 standard; Protein; 118 AA.

AC R63158;  
 DT 26-MAY-1995 (first entry)  
 DE Tumour rejection antigen precursor.  
 KW Tumour rejection antigen; precursor; HLA-A2 molecule; tyrosinase;  
 KW isolation; melanoma; cell line; LB-39-MEL; diagnosis; vaccine;  
 KW therapy.  
 OS Homo sapiens.  
 PN WO9529193-A2.  
 PD 28-SEP-1994.  
 PF 09-MAR-1994; U02487.  
 PR 18-MAR-1993; US-032978.  
 PA (LUDW-) LUDWIG INST CANCER RES.  
 PI Boon-Falleur T, Brichard V, De Plaen E, Traversari C;  
 PI Van Pei A, Wolfel T;  
 DR WPI; 94-316544/39.  
 DR N-PSDB; 076370.  
 PT Nucleic acid coding for a tumour rejection antigen precursor - is  
 PT used for developing prods. for diagnosis or treatment of expression  
 PT related disorders, partic. melanoma  
 PS Claim 5; Page 14; 26pp; English.  
 CC This sequence represents the tumour rejection antigen precursor which is  
 CC processed to a tumour rejection antigen presented by HLA-A2 molecules.  
 CC The tumour rejection antigen is not related to tyrosinase. The cDNA  
 CC encoding this sequence was isolated from the melanoma cell line,  
 CC LB-39-MEL. The tumour rejection antigen may be used for diagnosis or  
 CC in vaccines or for therapy of disorders characterised by the expression  
 CC of the tumour rejection antigen precursor, particularly melanoma.  
 SQ Sequence 118 AA;

Query Match 100.0%; Score 62; DB 12; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 1.03e+00;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 eaagigiltv 35  
 QY 1 EAAGIGILTV 10

RESULT 3  
 ID R84212 standard; Protein; 118 AA.  
 AC R84212; (first entry)  
 DT 20-APR-1996  
 DE MART-1 melanoma antigen.  
 KW MART-1; melanoma antigen recognised by T-cell; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian.  
 PN Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DR N-PSDB; T02714.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 11; Page 17; 184pp; English.  
 CC The melanoma antigen (MART-1) is produced by recombinant DNA  
 CC methods, i.e. preferably using a baculovirus vector for expression  
 CC in insect cell cultures. MART-1 protein is a source of immunogenic  
 CC peptides (see R84196 for peptide M9-2) which are optionally modified  
 CC (see R84783-R84800) and used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 118 AA;

Query Match 100.0%; Score 62; DB 12; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 1.03e+00;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 eaagigiltv 35  
 QY 1 EAAGIGILTV 10

RESULT 3  
 ID R84212 standard; Protein; 118 AA.  
 AC R84212; (first entry)  
 DT 20-APR-1996  
 DE MART-1 melanoma antigen.  
 KW MART-1; melanoma antigen recognised by T-cell; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian.  
 PN Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DR N-PSDB; T02714.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 11; Page 17; 184pp; English.  
 CC The melanoma antigen (MART-1) is produced by recombinant DNA  
 CC methods, i.e. preferably using a baculovirus vector for expression  
 CC in insect cell cultures. MART-1 protein is a source of immunogenic  
 CC peptides (see R84196 for peptide M9-2) which are optionally modified  
 CC (see R84783-R84800) and used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 118 AA;

Query Match 100.0%; Score 62; DB 15; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 1.03e+00;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 eaagigiltv 35  
 QY 1 EAAGIGILTV 10

RESULT 4  
 ID R84196 standard; Peptide; 9 AA.  
 AC R84196;  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen immunogenic peptide M9-2.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 12; Page 117; 184pp; English.  
 CC Immunogenic peptide M9-2 is based on the melanoma antigen (MART-1)  
 CC (see R84212). M9-2 may be modified to improve immunogenicity or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 90.3%; Score 56; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 5.22e+00;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aagigiltv 9  
 QY 2 AAGIGILTV 10

RESULT 5  
 ID R84198 standard; Peptide; 10 AA.  
 AC R84198;  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen immunogenic peptide M10-4.  
 KW MART-1; M10-4; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 12; Page 122; 184pp; English.  
 CC Immunogenic peptide M10-4 is a derivative of peptide M9-2 (R84196)

CC which is based on the melanoma antigen (MART-1) (see R84212).  
 CC M9-2 may be modified to improve immunogenicity (see R84783-R84800).  
 CC and used in medicaments for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against MART-1 and its  
 CC immunogenic peptides may be used in the detection and isolation of  
 CC MART-1 from a sample, the detection of which is indicative of a  
 CC disease state (melanoma or metastatic melanoma).  
 CC See also R84196.  
 SQ Sequence 10 AA;

Query Match 90.3%; Score 56; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 5.22e+00; Mismatches 0; Indels 0; Gaps 0;  
 Matches 9; Conservative 0;

Db 1 aagigiltv 9  
 |||||  
 QY 2 AAGIGILTV 10

RESULT 6  
 ID R84788 standard; Peptide; 9 AA.  
 AC R84788;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1Y.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 PI WPI; 95-382963/49.  
 DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 83.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.50e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||  
 QY 3 AGIGILTV 10

RESULT 7  
 ID R84787 standard; Peptide; 9 AA.  
 AC R84787;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1F.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.

PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 PI WPI; 95-382963/49.  
 DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 83.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.50e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||  
 QY 3 AGIGILTV 10

RESULT 8  
 ID R84786 standard; Peptide; 9 AA.  
 AC R84786;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1W.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 PI WPI; 95-382963/49.  
 DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 83.9%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.50e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||  
 QY 3 AGIGILTV 10

RESULT 9  
 ID R84785 standard; Peptide; 9 AA.  
 AC R84785;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2I.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;



KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 82.3%; Score 51; DB 15; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.95e+01;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 aligilgtv 9  
 | | | | | | |  
 QY 2 AAGIGILTV 10

RESULT 10  
 ID R84784 standard; Peptide; 9 AA.  
 AC R84784;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2M.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 80.6%; Score 50; DB 15; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 2.53e+01;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 amigilgtv 9  
 | | | | | | |  
 QY 2 AAGIGILTV 10

RESULT 11  
 ID R84783 standard; Peptide; 9 AA.  
 AC R84783;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2L.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 79.0%; Score 49; DB 15; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.27e+01;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 aligilgtv 9  
 | | | | | | |  
 QY 2 AAGIGILTV 10

RESULT 12  
 ID R84792 standard; Peptide; 9 AA.  
 AC R84792;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1K2L.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 77.4%; Score 48; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 4.23e+01;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigitlv 9  
|||||  
Qy 4 GIGILTV 10

RESULT 13  
ID R84764 standard; Peptide; 9 AA.

AC R84764;  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen immunogenic peptide M9-3.  
KW MART-1; M9-3; melanoma antigen recognised by T-cells; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.

OS Synthetic.  
PN WO9529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Example 2; Page 72; 184pp; English.  
CC Immunogenic peptide M9-3 is based on the melanoma antigen (MART-1)  
CC (see R84764). M9-3 may be modified to improve immunogenicity  
CC (see R84763-R84782) and used in medicaments (vaccines) for  
CC treatment or prevention (by immunization) of melanoma.  
CC Antibodies against MART-1 and its immunogenic peptides may be used  
CC in the detection and isolation of MART-1 from a sample, the  
CC detection of which is indicative of a disease state (melanoma  
CC or metastatic melanoma).  
CC See also R84196-R84198.  
SQ Sequence 9 AA;

Query Match 77.4%; Score 48; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.23e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 gigitlv 7  
|||||  
Qy 4 GIGILTV 10

RESULT 14  
ID R84796 standard; Peptide; 9 AA.

AC R84796;  
DT 25-APR-1996 (first entry)  
DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1F2L.  
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.

OS Synthetic.  
PN WO9529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Example 5; Page 105; 184pp; English.  
CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
CC an immunogenic peptide based on the melanoma antigen (MART-1) (see

CC R84212). The peptides are used in medicaments for the treatment or  
CC prevention (by immunization) of melanoma. Antibodies against MART-1  
CC and its immunogenic peptides may be used in the detection and  
CC isolation of MART-1 from a sample, the detection of which is  
CC indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 77.4%; Score 48; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.23e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigitlv 9  
|||||  
Qy 4 GIGILTV 10

RESULT 15

ID R84795 standard; Peptide; 9 AA.  
AC R84795;  
DT 25-APR-1996 (first entry)  
DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1W2L.  
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN WO9529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Example 5; Page 105; 184pp; English.  
CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
CC R84212). The peptides are used in medicaments for the treatment or  
CC prevention (by immunization) of melanoma. Antibodies against MART-1  
CC and its immunogenic peptides may be used in the detection and  
CC isolation of MART-1 from a sample, the detection of which is  
CC indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 77.4%; Score 48; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.23e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigitlv 9  
|||||  
Qy 4 GIGILTV 10

Search completed: Tue Jun 10 11:19:29 1997  
Job time : 8 secs.

\*\*\*\*\*  
WIREH (TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:18:56 1997; MasPar time 2.58 Seconds  
Tabular output not generated. 110.579 Million cell updates/sec

Title: >US-08-231-565A-17  
Description: (1-10) from US08231565A.pep  
Perfect Score: 62  
Sequence: 1 EAAGIGILTV 10

Scoring table: PAM 150  
Gap 15

Searched: 89912 seqs, 28507787 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir50  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 22.055; Variance 28.889; scale 0.763

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	62	100.0	118	13	A55253 melanoma antigen MAR	9.45e-03
2	62	100.0	118	13	I38506 melan-A protein - hu	9.45e-03
3	48	77.4	333	1	DEALXE acetoin[2,6-dichloro	8.29e+00
4	47	75.8	101	10	A33351 H+-transporting ATP	1.29e+01
5	47	75.8	110	4	VRBPDL major capsid protein	1.29e+01
6	47	75.8	479	4	VGBEPB glycoprotein gIII pr	1.29e+01
7	47	75.8	753	1	A42863 5-methyltetrahydropt	1.29e+01
8	46	74.2	231	12	S48276 YSA1 protein - yeast	2.00e+01
9	46	74.2	271	11	S38115 hypothetical protein	2.00e+01
10	45	72.6	339	16	S62369 methylcobalamin: Coe	3.08e+01
11	45	72.6	420	14	S59131 Kan-1 protein - rat	3.08e+01
12	45	72.6	744	14	A43353 ascites sialoglycopr	3.08e+01
13	44	71.0	635	10	S19011 endo-1,4-beta-xylana	4.71e+01
14	43	69.4	132	5	S01903 H+-transporting ATP	7.14e+01
15	43	69.4	345	1	KZSC alkaline phosphatase	7.14e+01
16	43	69.4	493	3	ACMSE nicotinic acetylchol	7.14e+01
17	43	69.4	590	11	A45772 nitrate-inducible ni	7.14e+01
-18	43	69.4	611	8	JT0592 hypothetical protein	7.14e+01
19	43	69.4	675	12	S53832 NADH dehydrogenase c	7.14e+01
20	43	69.4	1217	2	EGMSMG epidermal growth fac	7.14e+01
21	43	69.4	1436	11	S67655 probable membrane pr	7.14e+01

22 42 67.7 19 10 A34524 31K flagellin - Meth 1.08e+02  
23 42 67.7 216 10 C41316 flagellin B2 precurs 1.08e+02  
24 42 67.7 218 10 B41316 flagellin B1 precurs 1.08e+02  
25 42 67.7 251 9 S56416 hypothetical protein 1.08e+02  
26 42 67.7 264 9 PS0398 hypothetical protein 1.08e+02  
27 42 67.7 302 9 A55214 carboxysome formatio 1.08e+02  
28 42 67.7 308 9 H64153 hypothetical protein 1.08e+02  
29 42 67.7 481 9 C64107 thiophene and furan 1.08e+02  
30 42 67.7 501 10 S18573 L-lysine transport p 1.08e+02  
31 42 67.7 530 11 S61034 AIP2 protein - yeast 1.08e+02  
32 42 67.7 661 10 C36912 NADH-flavin oxidored 1.08e+02  
33 42 67.7 776 11 S45495 isp4 protein - fissi 1.08e+02  
34 42 67.7 1530 11 S52239 hba2 protein - fissi 1.08e+02  
35 41 66.1 226 3 JX0221 CD9 antigen - bovine 1.61e+02  
36 41 66.1 237 10 F64506 2-hydroxyhepta-2,4-d 1.61e+02  
37 41 66.1 323 9 A64054 selenium metabolism 1.61e+02  
38 41 66.1 387 11 S30831 GLY1 protein - yeast 1.61e+02  
39 41 66.1 395 13 I37165 ALI-1 protein - huma 1.61e+02  
40 41 66.1 401 16 S62465 hypothetical protein 1.61e+02  
41 41 66.1 734 7 S00704 photosystem I protei 1.61e+02  
42 41 66.1 845 14 A40016 matrix 3 - rat 1.61e+02  
43 41 66.1 1464 13 S47555 N-methyl-D-aspartate 1.61e+02  
44 41 66.1 3869 14 A48205 All-1 protein +cTE f 1.61e+02  
45 41 66.1 3988 13 A44265 tri thorax homolog HT 1.61e+02

ALIGNMENTS

RESULT 1  
ENTRY A55253 #type complete  
TITLE melanoma antigen MART-1 - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 06-Feb-1995  
ACCESSIONS A55253  
REFERENCE A55253  
#authors Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Rivoltini, L.; Topalian, S.L.; Miki, T.; Rosenberg, S.A.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1994) 91:3515-3519  
#title Cloning of the gene coding for a shared human melanoma antigen recognized by autologous T cells infiltrating into tumor.  
#accession A55253  
##status preliminary  
##molecule\_type mRNA  
##residues 1-118 #label KAW  
##cross-references GB:006452  
SUMMARY #length 118 #molecular-weight 13157 #checksum 3535

Query Match 100.0%; Score 62; DB 13; Length 118;  
Best Local Similarity 100.0%; Pred. No. 9.45e-03;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db -26 eaagigilv 35  
QY 1 EAAGIGILTV 10  
RESULT 2  
ENTRY I38506 #type complete  
TITLE melan-A protein - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 06-Sep-1996  
ACCESSIONS I38506  
REFERENCE I38506  
#authors Coulic, P.G.; Brichard, V.; Van Pel, A.; Wolfel, T.; Schneider, J.; Traversari, C.; Mattei, S.; De Plaen, E.; Lurquin, C.; Szikora, J.P.; Renaud, J.; Boon, T.  
#journal J. Exp. Med. (1994) 180:35-42  
#title A new gene coding for a differentiation antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas [see comments].

```

#cross-references MUID:94275389
#accession I38506
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-118 ##label RES
##cross-references EMBL:U06654; NID:g517022; CDS_PID:g517023
SUMMARY #length 118 #molecular-weight 13157 #checksum 3535

Query Match 100.0%; Score 62; DB 13; Length 118;
Best Local Similarity 100.0%; Pred. No. 9.45e+03;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 eaagigiltv 35
|||||
QY 1 EAAGIGILTV 10

RESULT 3
ENTRY DEALXE #type complete
TITLE acetoin[2,6-dichlorophenolindophenol] oxidoreductase (EC
1.-.-) alpha chain - Alcaligenes eutrophus (strain H16)
ORGANISM #formal_name Alcaligenes eutrophus
DATE 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change
27-Jan-1995
ACCESSIONS B42462
REFERENCE A42462
#authors Priefert, H.; Hein, S.; Krueger, N.; Zeh, K.; Schmidt, B.;
Steinbuechel, A.
#journal J. Bacteriol. (1991) 173:4056-4071
#title Identification and molecular characterization of the
Alcaligenes eutrophus H16 aco operon genes involved in
acetoin catabolism.
#cross-references MUID:91286190
#accession B42462
##molecule_type DNA
##residues 1-333 ##label PRI
##cross-references GB:M66060
COMMENT This is a component of the enzyme complex that catalyzes 2,
6-dichlorophenolindophenol-dependent cleavage of acetoin into
acetate and acetaldehyde. The functional enzyme is a tetramer of
two alpha and two beta chains.
CLASSIFICATION #superfamily pyruvate dehydrogenase (lipoamide) alpha chain;
thiamine pyrophosphate-binding domain homology
KEYWORDS heterotetramer; oxidoreductase
FEATURE
145-194 #domain thiamine pyrophosphate-binding domain homology
#label TPB
SUMMARY #length 333 #molecular-weight 35375 #checksum 2647

Query Match 77.4%; Score 48; DB 1; Length 333;
Best Local Similarity 87.5%; Pred. No. 8.29e+00;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 56 eaagvgil 63
|||||
QY 1 EAAGIGIL 8

RESULT 4
ENTRY #type complete
TITLE H+-transporting ATP synthase (EC 3.6.1.34) proteolipid chain
- Sulfolobus acidocaldarius
ORGANISM #formal_name Sulfolobus acidocaldarius
DATE 20-Dec-1989 #sequence_revision 20-Dec-1989 #text_change
23-Jun-1993
ACCESSIONS A33351
REFERENCE A33351
#authors Denda, K.; Konishi, J.; Oshima, T.; Date, T.; Yoshida, M.
#journal J. Biol. Chem. (1989) 264:7119-7121
#title A gene encoding the proteolipid subunit of Sulfolobus
acidocaldarius ATPase complex.
#cross-references MUID:89214142
#accession A33351

##status preliminary
##molecule_type DNA
##residues 1-101 ##label DEN
##cross-references GB:J04740
KEYWORDS hydrolase
SUMMARY #length 101 #molecular-weight 10362 #checksum 4300

Query Match 75.8%; Score 47; DB 10; Length 101;
Best Local Similarity 87.5%; Pred. No. 1.29e+01;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 59 aagigiltv 66
|||||
QY 2 AAGIGILT 9

RESULT 5
ENTRY VHBPDL #type complete
TITLE major capsid protein D - phage lambda
ALTERNATE_NAMES head protein D
ORGANISM #formal_name phage lambda
DATE 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change
17-Nov-1995
ACCESSIONS G04333; C43013; A04334; A23206
REFERENCE A94614
#authors Daniels, D.
#submision submitted to the Nucleic Acid Sequence Database, September
1982
#accession G04333
##molecule_type DNA
##residues 1-110 ##label DAN
REFERENCE A92891
#authors Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen,
G.B.
#journal J. Mol. Biol. (1982) 162:729-773
#title Nucleotide sequence of bacteriophage lambda DNA.
#cross-references MUID:83189071
#accession C43013
##molecule_type DNA
##residues 1-110 ##label SAN
REFERENCE A23206
#authors Witkiewicz, H.; Schweiger, M.
#journal EMBO J. (1982) 1:1559-1564
#cross-references MUID:84207913
#contents annotation; physicochemical properties
COMMENT Gene D protein is a major component of the phage head and serves to
stabilize the head during DNA packaging. There are approximately
420 copies of protein D per mature phage.
GENETICS
#gene D
#map_position 11.85-12.53
CLASSIFICATION #superfamily phage lambda major capsid protein D
KEYWORDS DNA packaging
SUMMARY #length 110 #molecular-weight 11572 #checksum 3863

Query Match 75.8%; Score 47; DB 4; Length 110;
Best Local Similarity 50.0%; Pred. No. 1.29e+01;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 55 dgaavgilav 64
|||||
QY 1 EAAGIGILTV 10

RESULT 6
ENTRY VGBEPB #type complete
TITLE glycoprotein gIII precursor - suid herpesvirus 1
ORGANISM #formal_name suid herpesvirus 1
DATE 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change
05-Jan-1996
ACCESSIONS A26097
REFERENCE A26097
#authors Robbins, A.K.; Watson, R.J.; Whealy, M.E.; Hays, W.W.;

```

```

Enquist, L.W.
#journal J. Virol. (1986) 58:339-347
#title Characterization of a Pseudorabies virus glycoprotein gene
        with homology to herpes simplex virus type 1 and type 2
        glycoprotein C.
#cross-references MUID:86200375
#accession A36097
#molecule_type DNA
##residues 1-479 ##label ROB
##experimental_source strain Becker
CLASSIFICATION #superfamily herpesvirus glycoprotein F
KEYWORDS glycoprotein
FEATURE
1-22 #domain signal sequence #status predicted #label SIG\
23-479 #product glycoprotein gIII #status predicted #label GPG\
40,84,169,192,220,
228,285,302
SUMMARY #length 479 #molecular-weight 51206 #checksum 1630
        Predicted
Query Match 75.8%; Score 47; DB 4; Length 479;
Best Local Similarity 75.0%; Pred. No. 1.29e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigllai 463
QY 3 AGIGILTV 10

RESULT 7
ENTRY #type complete
TITLE 5-methyltetrahydropteroyltriglutamate-homocysteine
        S-methyltransferase (EC 2.1.1.14) - Escherichia coli
ALTERNATE_NAMES cobalamin-independent methionine synthase;
        tetrahydropteroyltriglutamate methyltransferase
ORGANISM #formal_name Escherichia coli
DATE 17-Feb-1994 #sequence_revision 06-Sep-1996 #text_change
        06-Sep-1996
ACCESSIONS A42863; S30719
REFERENCE A42863
#authors Gonzalez, J.C.; Banerjee, R.V.; Huang, S.; Summer, J.S.;
        Matthews, R.G.
#journal Biochemistry (1992) 31:6045-6056
#title Comparison of cobalamin-independent and cobalamin-dependent
        methionine synthases from Escherichia coli: two solutions
        to the same chemical problem.
#cross-references MUID:92329421
#accession A42863
#molecule_type DNA
##residues 1-753 ##label GON
##cross-references NCBI:P109176
##experimental_source strain DH5alphaF'
##note sequence extracted from NCBI backbone
REFERENCE S30660
#authors Daniels, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.
#journal Science (1992) 257:771-778
#title Analysis of the Escherichia coli genome: DNA sequence of the
        region from 84.5 to 86.5 minutes.
#accession S30719
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
##residues 1-362,'L',364-580,'CYMKWPWIKPFLASRLNRR',603-604,'Q',
        606-638,'R',660-753 ##label DAN
##cross-references EMBL:M87049
##note the nucleotide sequence was submitted to the EMBL Data
        Library, November 1992
##note the sequence of residues 581-602 appears to be in error
        by homology with PIR2:B64137, PIR2:S50594, and
        PIR2:S57636

GENETICS metE
#gene #superfamily cobalamin-independent methionine synthase
CLASSIFICATION duplication; methionine biosynthesis; methyltransferase
KEYWORDS

```

```

FEATURE
726 #active_site Cys (methylcysteine intermediate) #status
        experimental
SUMMARY #length 753 #molecular-weight 84659 #checksum 6645
Query Match 75.8%; Score 47; DB 1; Length 753;
Best Local Similarity 70.0%; Pred. No. 1.29e+01;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 589 eaagigilqi 598
QY 1 EAAGIGILTV 10

RESULT 8
ENTRY #type complete
TITLE YSA1 protein - yeast (Saccharomyces cerevisiae)
        protein YBR0907; protein YBR111c
ALTERNATE_NAMES #formal_name Saccharomyces cerevisiae
ORGANISM 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change
        01-Sep-1995
ACCESSIONS S48276; S45979; S25364; S44691
REFERENCE S48255
#authors Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann,
        H.
#journal Yeast (1994) 10:1363-1381
#title Analysis of a 70 kb region on the right arm of yeast
        chromosome II.
#accession S48276
#molecule_type DNA
##residues 1-231 ##label MAN
##cross-references EMBL:X78993
##note nucleotide sequence is not given
REFERENCE S45927
#authors Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
#submission submitted to the Protein Sequence Database, August 1994
#accession S45979
#molecule_type DNA
##residues 1-231 ##label FE2
##cross-references EMBL:Z35980
REFERENCE S25364
#authors Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann,
        H.
#journal Yeast (1992) 8:397-408
#title Molecular analysis of yeast chromosome II between CMD1 and
        YFS2: the excision repair gene RAD16 located in this region
        belongs to a novel group of double-finger proteins.
#accession S25364
#molecule_type DNA
##residues 1-47 ##label MAW
##cross-references EMBL:X66247
GENETICS LISTA:YSA1
#gene #map_position 2R
SUMMARY #length 231 #molecular-weight 26087 #checksum 4809
Query Match 74.2%; Score 46; DB 12; Length 231;
Best Local Similarity 85.7%; Pred. No. 2.00e+01;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 79 qigilti 85
QY 4 GIGILTV 10

RESULT 9
ENTRY #type complete
TITLE hypothetical protein YKR043c - yeast (Saccharomyces
        cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 03-May-1994 #sequence_revision 03-May-1994 #text_change
        12-Apr-1996
ACCESSIONS S38115

```

```

REFERENCE      S38097
#authors       Urrestarazu, L.A.; Jauniaux, J.C.
#submission    submitted to the Protein Sequence Database, March 1994
#accession     S38115
#molecule_type DNA
#residues      1-271 ##label URR
##cross-references EMBL:Z28268
##experimental_source strain S288C
GENETICS
#map_position  11R
SUMMARY        #length 271 #molecular-weight 31022 #checksum 8533
Query Match    74.2%; Score 46; DB 11; Length 271;
Best Local Similarity 55.6%; Pred. No. 2.00e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 233 dagglgvlvs 241
:|:|:|:|
QY 1 EAAGIGILT 9

RESULT 10
ENTRY      S62369 #type complete
TITLE      methylobalamin: Coenzyme M methyltransferase (isoenzyme II)
ORGANISM   #formal_name Methanosarcina barkeri
DATE       20-Jul-1996 #sequence_revision 20-Jul-1996 #text_change
ACCESSIONS S62369
REFERENCE   S62368
#authors    Harns, U.; Thauer, R.K.
#journal     Eur. J. Biochem. (1996) 235:653-659
#title       Methylobalamin:coenzyme M methyltransferase isoenzymes MtaA
              and MtaB from Methanosarcina barkeri. Cloning, sequencing
              and differential transcription of the encoding genes, and
              functional overexpression of the mtaA gene in Escherichia
              coli.
#accession   S62369
##status     preliminary
##residues   1-339 ##label HAR
##cross-references EMBL:X91894
SUMMARY      #length 339 #molecular-weight 36761 #checksum 6431

Query Match    72.6%; Score 45; DB 16; Length 339;
Best Local Similarity 75.0%; Pred. No. 3.08e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 307 agvglltv 314
:|:|:|:|
QY 3 AGIGILT 10

RESULT 11
ENTRY      S59131 #type complete
TITLE      Kan-1 protein - rat
ORGANISM   #formal_name Rattus norvegicus #common_name Norway rat
DATE       15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
ACCESSIONS S59131
REFERENCE   S59131
#authors    Furutani, M.; Arai, S.; Higashitsuji, H.; Mise, M.; Fukumoto,
              M.; Takano, S.; Nakayama, H.; Imamura, M.; Fujita, J.
#journal     Biochem. J. (1995) 311:203-208
#title       Reduced expression of kan-1 (encoding putative bile
              acid-CoA-amino acid N-acyltransferase) mRNA in livers of
              rats after partial hepatectomy and during sepsis.
#accession   S59131
##status     preliminary
##molecule_type mRNA
##residues   1-420 ##label FUR
##cross-references EMBL:DA3964
SUMMARY      #length 420 #molecular-weight 46496 #checksum 4868

REFERENCE      S38097
#authors       Urrestarazu, L.A.; Jauniaux, J.C.
#submission    submitted to the Protein Sequence Database, March 1994
#accession     S38115
#molecule_type DNA
#residues      1-271 ##label URR
##cross-references EMBL:Z28268
##experimental_source strain S288C
GENETICS
#map_position  11R
SUMMARY        #length 271 #molecular-weight 31022 #checksum 8533
Query Match    74.2%; Score 46; DB 11; Length 271;
Best Local Similarity 55.6%; Pred. No. 2.00e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 233 dagglgvlvs 241
:|:|:|:|
QY 1 EAAGIGILT 9

RESULT 12
ENTRY      A43353 #type fragment
TITLE      ascites sialoglycoprotein-2 - rat (fragment)
ORGANISM   #formal_name Rattus norvegicus #common_name Norway rat
DATE       10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change
ACCESSIONS A43353
REFERENCE   A43353
#authors    Sheng, Z.; Wu, K.; Carraway, K.L.; Fregien, N.
#journal     J. Biol. Chem. (1992) 267:16341-16346
#title       Molecular cloning of the transmembrane component of the 13762
              mammary adenocarcinoma sialomucin complex. A new member of
              the epidermal growth factor superfamily.
#cross-references MUID:92355597
#accession   A43353
##status     preliminary
##molecule_type mRNA; protein
##residues   1-744 ##label SHE
##cross-references NCBI:110690; NCBI:P110691
##experimental_source mammary adenocarcinoma
##note       #length 744 #checksum 2462
SUMMARY      #length 744 #checksum 2462

Query Match    72.6%; Score 45; DB 14; Length 744;
Best Local Similarity 60.0%; Pred. No. 3.08e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 201 etnglgilgv 210
:|:|:|:|
QY 1 EAAGIGILT 10

RESULT 13
ENTRY      S19011 #type complete
TITLE      endo-1,4-beta-xylanase (EC 3.2.1.8) - Bacillus polymyxa
ORGANISM   #formal_name Bacillus polymyxa
DATE       22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
ACCESSIONS S19011
REFERENCE   S19011
#authors    Gosalbes, M.J.; Perez-Gonzalez, J.A.; Gonzalez, R.; Navarro,
              A.
#journal     J. Bacteriol. (1991) 173:7705-7710
#title       Two beta-glycanase genes are clustered in Bacillus polymyxa:
              molecular cloning, expression, and sequence analysis of
              genes encoding a xylanase and an endo-beta-(1,3)-{1,
              4)-glucanase.
#cross-references MUID:92041687
#accession   S19011
##status     preliminary
##molecule_type DNA
##residues   1-635 ##label GOS
##cross-references EMBL:X57094
##note       the authors translated the codon GAA for residue 78 as
              Gly, CCT for residue 272 as Thr, ATC for residue 412
              as Gln, and ATC for residue 478 as Tyr
KEYWORDS     glycosidase; hydrolase
SUMMARY      #length 635 #molecular-weight 67914 #checksum 2077

Query Match    71.0%; Score 44; DB 10; Length 635;
Best Local Similarity 75.0%; Pred. No. 4.71e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 149 gagigvlt 156

```

```
QY      2 AAGIGILT 9
      :||||:|
RESULT  14
ENTRY   S01903      #type complete
TITLE   H+-transporting ATP synthase (EC 3.6.1.34) epsilon chain -
ORGANISM Arabidopsis thaliana chloroplast
FORMAL_NAME chloroplast Arabidopsis thaliana #common_name
mouse-ear cress
DATE    01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change
14-Sep-1994
ACCESSIONS S01903
REFERENCE  Chen, H.C.; Wintz, H.; Weill, J.H.; Pillay, D.T.N.
Nucleic Acids Res. (1988) 16:10372
Nucleotide sequence of chloroplast CFI-ATPase epsilon-subunit
and elongator tRNA(Met) genes from Arabidopsis thaliana.
#cross-references MUID:89057486
#accession S01903
#molecule_type DNA
#residues 1-132 #label CHE
#cross-references EMBL:X12889
GENETICS
#gene atpE
#genome chloroplast
CLASSIFICATION #superfamily H+-transporting ATP synthase epsilon chain
KEYWORDS ATP synthesis; chloroplast; hydrolase; membrane-associated
complex; thylakoid
SUMMARY #length 132 #molecular-weight 14472 #checksum 1607
Query Match 69.4%; Score 43; DB 5; Length 132;
Best Local Similarity 66.7%; Pred. No. 7.14e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 43 avdigilti 51
| | | | |
QY 2 AAGIGILT 10

RESULT  15
ENTRY   KZEC      #type complete
TITLE   alkaline phosphatase isozyme conversion protein (EC 3.4.11.-)
- Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE    31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
23-Aug-1996
ACCESSIONS A28382
REFERENCE  A28382
AUTHORS  Ishino, Y.; Shinagawa, H.; Makino, K.; Amemura, M.; Nakata,
A.
J. Bacteriol. (1987) 169:5429-5433
Nucleotide sequence of the iap gene, responsible for alkaline
phosphatase isozyme conversion in Escherichia coli, and
identification of the gene product.
#cross-references MUID:88058748
#accession A28382
#molecule_type DNA
#residues 1-345 #label ISH
GENETICS
#gene iap
#map_position 59 min
FUNCTION
#description this protein, presumably an aminopeptidase, mediates the
conversion of E. coli alkaline phosphatase from isozyme 1
to isozymes 2 and 3 by proteolitically removing the two
amino-terminal arginine residues one by one
CLASSIFICATION #superfamily alkaline phosphatase isozyme conversion protein
KEYWORDS alpha-aminoacylpeptide hydrolase
FEATURE
1-24 #domain signal sequence #status predicted #label SIG\
25-345 #product alkaline phosphatase isozyme conversion protein
#status predicted #label MAT
```

```
SUMMARY      #length 345 #molecular-weight 37920 #checksum 1871
Query Match 69.4%; Score 43; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 7.14e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 267 dkagialvsv 276
: | | | | |
QY 1 EAAGIGILT 10
Search completed: Tue Jun 10 11:19:04 1997
Job time : 8 secs.
```

\*\*\*\*\*  
M I S E R I E (TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:18:30 1997; MasPar time 1.99 Seconds  
Tabular output not generated. 106.771 Million cell updates/sec

Title: >US-08-231-565A-17  
Description: (1-10) from US08231565A.pap  
Perfect Score: 62  
Sequence: 1 EAAGIGILTV 10

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 23.060; Variance 25.033; scale 0.921

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	48	77.4	332	1	ACOA_ALCEU ACETOIN:2,6-DICHLOROP	2.67e+00
2	47	75.8	101	1	ATPL_SULAC MEMBRANE-ASSOCIATED A	4.45e+00
3	47	75.8	110	10	VCAD_LAMBD HEAD DECORATION PROTE	4.45e+00
4	47	75.8	479	10	VGLC_PRVIF GLYCOPROTEIN GIII PRE	4.45e+00
5	47	75.8	752	6	METE_ECOLI 5-METHYLTETRAHYDROPE	4.45e+00
6	46	74.2	231	11	YSAL_YEAST YSAL PROTEIN.	7.36e+00
7	46	74.2	271	11	YK23_YEAST HYPOTHETICAL 31.0 KD	7.36e+00
8	44	71.0	456	4	GLMD_ECOLI UDP-N-ACETYLGLUCOSAM	1.96e+01
9	44	71.0	467	4	FLID_ECOLI FLAGELLAR HOOK-ASSOCI	1.96e+01
10	44	71.0	635	10	XYND_BACPO ENDO-1,4-BETA-XYLANAS	1.96e+01
11	43	69.4	132	1	ATPE_ARATH ATP SYNTHASE EPSILON	3.15e+01
12	43	69.4	207	8	RL4_BACSU 50S RIBOSOMAL PROTEIN	3.15e+01
13	43	69.4	345	5	IAP_ECOLI ALKALINE PHOSPHATASE	3.15e+01
14	43	69.4	388	11	YXBC_BACSU HYPOTHETICAL METABOLI	3.15e+01
15	43	69.4	493	1	ACHE_MOUSE ACETYLCHOLINE RECEPTO	3.15e+01
16	43	69.4	590	2	CHLL_ARATH NITRATE/CHLORATE TRAN	3.15e+01
17	43	69.4	611	11	YD3M_HERAU HYPOTHETICAL 68.4 KD	3.15e+01
18	43	69.4	1217	3	EGF_MOUSE EPIDERMAL GROWTH FACT	3.15e+01
19	42	67.7	216	4	FLA2_METVO FLAGELLIN B2 PRECURSO	5.02e+01
20	42	67.7	218	4	FLA1_METVO FLAGELLIN B1 PRECURSO	5.02e+01
21	42	67.7	251	11	YJFQ_ECOLI HYPOTHETICAL TRANSORI	5.02e+01
22	42	67.7	308	6	MENA_HAEIN 1,4-DIHYDROXY-2-NAPHT	5.02e+01

23	42	67.7	321	11	YHBE_ECOLI HYPOTHETICAL 35.0 KD	5.02e+01
24	42	67.7	461	9	THDF_HAEIN POSSIBLE THIOPHENE AN	5.02e+01
25	42	67.7	501	6	LYSI_CORGL L-LYSINE TRANSPORT PR	5.02e+01
26	42	67.7	530	1	ALP2_YEAST ACTIN INTERACTING PRO	5.02e+01
27	42	67.7	659	11	YVBT_BACSU HYPOTHETICAL 74.3 KD	5.02e+01
28	42	67.7	661	1	BAIH_EUBSP NADH-DEPENDENT FLAVIN	5.02e+01
29	42	67.7	776	5	ISP4_SCHPO SEXUAL DIFFERENTIATIO	5.02e+01
30	42	67.7	1530	1	BFR1_SCHPO BREFFELDIN A RESISTANC	5.02e+01
31	42	66.1	225	2	CD9_FELCA CD9 ANTIGEN.	7.92e+01
32	41	66.1	225	2	CD9_BOVIN CD9 ANTIGEN.	7.92e+01
33	41	66.1	348	9	SELD_HAEIN SELENIDE-WATER DIKINA	7.92e+01
34	41	66.1	359	11	YFDA_CORGL HYPOTHETICAL PROTEIN	7.92e+01
35	41	66.1	387	4	GLY1_YEAST GLY1 PROTEIN.	7.92e+01
36	41	66.1	467	4	HEM1_MYCLE GLUTAMYL-TRNA REDUCTA	7.92e+01
37	41	66.1	587	11	YIDU_ECOLI HYPOTHETICAL 64.0 KD	7.92e+01
38	41	66.1	650	10	Y00P_MYCTU PUTATIVE OXIDOREDUCTA	7.92e+01
39	41	66.1	666	2	CO14_BRAJA PROBABLE CYTOCHROME C	7.92e+01
40	41	66.1	734	7	PSAB_PEA PHOTOSYSTEM I P700 CH	7.92e+01
41	41	66.1	845	6	MAT3_RAT MATRIN 3.	7.92e+01
42	41	66.1	977	11	YDG8_SCHPO HYPOTHETICAL 111.4 KD	7.92e+01
43	41	66.1	3866	5	HRX_MOUSE ZINC FINGER PROTEIN H	7.92e+01
44	41	66.1	3969	5	HRX_HUMAN ZINC FINGER PROTEIN H	7.92e+01
45	41	66.1	5147	3	FAT_DROME CADHERIN-RELATED TUMO	7.92e+01

ALIGNMENTS

RESULT 1	ACOA_ALCEU	STANDARD;	PRT;	332 AA.
AC	P27745;			
DT	01-AUG-1992 (REL. 23, CREATED)			
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)			
DT	01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)			
DE	ACETOIN:2,6-DICHLOROPHENOLINDOPHENOL OXIDOREDUCTASE ALPHA			
DE	SUBUNIT (EC 1.1.1.-) (ACETOIN:DCPIP OXIDOREDUCTASE-ALPHA)			
DE	(AO:DCPIP OR).			
GN	ACOA.			
OS	ALCALIGENES EUTROPHUS.			
OC	PROKARYOTA: GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;			
OC	ALCALIGENACEAE.			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-31.			
RC	STRAIN=H16;			
RX	MEDLINE: 91286190.			
RA	PRIEFERT H., HEIN S., KRUEGER N., ZEH K., SCHMIDT B.,			
RA	STEINBUESCHEL A.;			
RL	J. BACTERIOL. 173:4056-4071(1991).			
CC	- - FUNCTION: CATALYZES THE 2,6-DICHLOROPHENOLINDOPHENOL-DEPENDENT CLEAVAGE OF ACETOIN INTO ACETATE AND ACETALDEHYDE, IN VITRO. THE ALPHA SUBUNIT IS PROBABLY THE CATALYTIC SUBUNIT OF THE ENZYME.			
CC	- - PATHWAY: ACETOIN CATABOLISM.			
CC	- - COFACTOR: THIAMINE PYROPHOSPHATE.			
CC	- - SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.			
CC	- - SIMILARITY: TO THE ALPHA SUBUNITS OF 2-OXO-ACID DEHYDROGENASE COMPONENTS OF VARIOUS MULTIENZYME COMPLEXES.			
DR	EMBL: M56060; G141894; -.			
DR	PIR: B42462; DEALAE.			
KW	OXIDOREDUCTASE; FLAVOPROTEIN; THIAMINE PYROPHOSPHATE.			
FT	INIT_MET 0			
FT	BINDING 173 173 THIAMINE PYROPHOSPHATE (BY SIMILARITY).			
SQ	SEQUENCE 332 AA; 35243 MW; 3322DA8E CRC32;			

Query Match 77.4%; Score 48; DB 1; Length 332;  
Best Local similarity 87.5%; Pred. No. 2.67e+00;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 55 eaagvgil 62

QY 1 EAAGIGIL 8

RESULT 2

ID ATPL\_SULAC STANDARD; PRT; 101 AA.



```

AC P23040;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE MEMBRANE-ASSOCIATED ATPASE C CHAIN (EC 3.6.1.34) (SUL-ATPASE
DE PROTEOLIPID CHAIN).
GN ATPP.
OS SULFOLOBUS ACIDOCALDARIUS.
RC ARCHAEABACTERIA; CRENARCHAEOTA; SULFOLOBALES.
RN [1]
RX MEDLINE; 89214142.
RA DENDA K., KONISHI J., OSHIMA T., DATE T., YOSHIDA M.;
RL J. BIOL. CHEM. 264:7119-7121(1989).
CC -1- FUNCTION: THE C CHAIN IS A PROTEOLIPID, AND ONE OF THE MEMBRANOUS
CC SUBUNITS OF THE NONENZYMATIC COMPONENT OF THE SUL-ATPASE
CC COMPLEX.
CC -1- SUBUNIT: SUL-ATPASE IS COMPOSED OF SIX (OR FIVE ?) SUBUNITS:
CC ALPHA, BETA, DELTA, GAMMA, C (PROTEOLIPID), AND POSSIBLY EPSILON.
CC -1- SIMILARITY: TO THE PROTEOLIPID SUBUNIT OF FO1F1-ATPASES.
DR EMBL; J04740; G152925; -.
DR PIR; A33351; A33351.
KW HYDROGEN ION TRANSPORT; LIPID-BINDING; TRANSMEMBRANE.
FT TRANSMEM 5 25 POTENTIAL.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
SQ SEQUENCE 101 AA; 10362 MW; 1DC8C74D CRC32;

Query Match 75.8%; Score 47; DB 1; Length 101;
Best Local Similarity 87.5%; Pred. No. 4.45e+00;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 59 aagigvlt 66
|||||:|
QY 2 AAGIGILT 9

RESULT 3
ID VCAD.LAMBED STANDARD; PRT; 110 AA.
AC P03712;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE HEAD DECORATION PROTEIN (GPD) (MAJOR CAPSID PROTEIN D).
GN D.
OS BACTERIOPHAGE LAMBDA.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; SIPHOVIRIDAE.
RN [1]
RX MEDLINE; 84207913.
RA WITKIEWICZ H., SCHWEIGER M.;
RL EMBL J. 1:1559-1564(1982).
CC -1- FUNCTION: STABILIZES THE HEAD SHELL FOLLOWING THE REARRANGEMENT
CC OF THE GPE SUBUNITS OF THE HEAD SHELL LATTICE THAT ACCOMPANIES
CC EXPANSION OF THE HEAD. THERE ARE APPROXIMATELY 420 COPIES OF
CC PROTEIN D PER MATURE PHAGE.
CC -1- SIMILARITY: TO BACTERIOPHAGE 21 HEAD DECORATION PROTEIN.
DR EMBL; J02459; G21511; -.
DR PIR; A04334; VHBPD.
DR PIR; A23206; A23206.
KW COAT PROTEIN.
SQ SEQUENCE 110 AA; 11572 MW; FDD50011 CRC32;

Query Match 75.8%; Score 47; DB 10; Length 110;
Best Local Similarity 50.0%; Pred. No. 4.45e+00;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 55 dgaavgilav 64

```

```

QY 1 EAAGIGILT 10
:::|:|:|:|
RESULT 4
ID VGLC_PRVIF STANDARD; PRT; 479 AA.
AC P06024;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE GLYCOPROTEIN GIII PRECURSOR.
OS PSEUDORABIES VIRUS (STRAIN INDIANA-FUNKHAUSER / BECKER) (PRV).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAPERPEVIRINAE.
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE; 86200375.
RA ROBBINS A.K., WATSON R.J., WHEALY M.E., HAYS W.W., ENQUIST L.W.;
RL J. VIROL. 58:339-347(1986).
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN C.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN GENE SUPERFAMILY.
DR EMBL; M12778; G334050; -.
DR PIR; A26097; VGBEPB.
KW GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.
FT SIGNAL 1 22
FT CHAIN 23 479 GLYCOPROTEIN GIII.
FT CARBOHYD 40 40 POTENTIAL.
FT CARBOHYD 84 84 POTENTIAL.
FT CARBOHYD 169 169 POTENTIAL.
FT CARBOHYD 192 192 POTENTIAL.
FT CARBOHYD 220 220 POTENTIAL.
FT CARBOHYD 228 228 POTENTIAL.
FT CARBOHYD 285 285 POTENTIAL.
FT CARBOHYD 302 302 POTENTIAL.
SQ SEQUENCE 479 AA; 51206 MW; 42EE5703 CRC32;

Query Match 75.8%; Score 47; DB 10; Length 479;
Best Local Similarity 75.0%; Pred. No. 4.45e+00;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigilal 463
|||||:|
QY 3 AGIGILT 10

RESULT 5
ID METE_ECOLI STANDARD; PRT; 752 AA.
AC P25665;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE 5-METHYLTRAHYDROPTEROYLTRIGLUTAMATE--HOMOCYSTEINE METHYLTRANSFERASE
DE (EC 2.1.1.14) (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME).
GN METE.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RX SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA MEDLINE; 92358234.
RA DANIELS D.L., PLUNKETT G. III, BURLAND V.D., BLATTNER F.R.;
RL SCIENCE 257:771-778(1992).
RN [2]
RX REVISIONS, SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA MEDLINE; 93347969.
RA PLUNKETT G. III, BURLAND V., DANIELS D.L., BLATTNER F.R.;
RL NUCLEIC ACIDS RES. 21:3391-3398(1993).
RN [3]
RX SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-K12 / DH5-ALPHA;
RX MEDLINE; 92329421.
RA GONZALEZ J.C., BANERJEE R.V., HUANG S., SUMNER J.S., MATTHEWS R.G.;

```

```
RL BIOCHEMISTRY 31:6045-6056(1992).
RN [4]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE; 89098936.
RA MAXON M.E., REDFIELD B., CAI X.-Y., SHOEMAN R., FUJITA K., FISHER W.,
RA STAUFFER G., WEISSBACH H., BROT N.;
RA PROC. NATL. ACAD. SCI. U.S.A. 86:85-89(1989).
RN [5]
RP SEQUENCE OF 1-12.
RC STRAIN-K12 / EMG2;
RA LINK A.J.;
RL SUBMITTED (OCT-1994) TO THE SWISS-PROT DATA BANK.
CC -1- CATALYTIC ACTIVITY: 5-METHYLTRAHYDROXYLARI-L-GLUTAMATE +
CC L-HOMOCYSTEINE - TETRAHYDROXYLARI-L-GLUTAMATE + L-METHIONINE.
CC -1- HAS AN ABSOLUTE REQUIREMENT FOR A POLYGLUTAMYLATED FOLATE AS
CC SUBSTRATE. ITS ACTIVITY DEPENDS ON PHOSPHATE ANIONS AND DIVALENT
CC CATIONS.
CC -1- SUBUNIT: MONOMER.
CC -1- PATHWAY: TERMINAL STEP IN THE DE NOVO BIOSYNTHESIS OF METHIONINE.
CC -1- SIMILARITY: TO THE EQUIVALENT ENZYME IN OTHER BACTERIAL AND
CC EUKARYOTIC SPECIES.
DR EMBL; M87049; G836660; -.
DR EMBL; M87625; G145474; -.
DR EMBL; J04155; G146827; -.
DR PIR; S30719; S30719.
DR PIR; A42863; A42863.
DR ECO2DBASE; F088.0; 6TH EDITION.
DR ECGENE; EG10584; METE.
DR TRANSFERASE; METHYLTRANSFERASE; METHIONINE BIOSYNTHESIS; REPEAT.
KW INT_MET 0
FT REPEAT 1 369 APPROXIMATE.
FT REPEAT 370 752 APPROXIMATE.
FT SITE 725 725 ITS ALKYLATION RESULTS IN COMPLETE
FT LOSS OF ACTIVITY.
FT CONFLICT 362 362 L -> V (IN REF. 3).
FT CONFLICT 604 604 Q -> E (IN REF. 3).
FT CONFLICT 658 658 R -> A (IN REF. 3).
SQ SEQUENCE 752 AA; 84626 MW; 55164250 CRC32;

Query Match 75.8%; Score 47; DB 6; Length 752;
Best Local Similarity 70.0%; Pred. No. 4.45e+00;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 588 eaagigilqi 597
QY 1 EAAGIGILT 10

RESULT 6
ID YSAL_YEAST STANDARD; PRT; 231 AA.
AC Q01976;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE YSAL PROTEIN.
GN YSAL OR YER111C OR YER0907.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE; 95208357.
RA MANNAHAPT G., STUCKA R., EHNLE S., VETTER I., FELDMANN H.;
RL YEAST 10:1363-1381(1994).
RN [2]
RP SEQUENCE OF 1-47 FROM N.A.
RC STRAIN-S288C;
RX MEDLINE; 92327848.
RA MANNAHAPT G., STUCKA R., EHNLE S., VETTER I., FELDMANN H.;
RL YEAST 8:397-408(1992).
CC -1- SIMILARITY: STRONG, TO B SUBTILIS YOKG.
CC -1- SIMILARITY: TO PROTEINS WITH A CORE MUTT DOMAIN.
DR EMBL; Z35980; G536466; -.

DR EMBL; X78993; G476067; -.
DR EMBL; X66247; G3549; -.
DR PIR; S44691; S44691.
DR LISTA; SC01415; YSAL.
DR SGD; L0002551; YSAL.
DR PROSITE; PS00893; MUTT.
FT DOMAIN 112 145 MUTT-LIKE.
SQ SEQUENCE 231 AA; 26087 MW; 49A2D6CB CRC32;

Query Match 74.2%; Score 46; DB 11; Length 231;
Best Local Similarity 85.7%; Pred. No. 7.36e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 79 g1gilti 85
QY 4 GIGILT 10

RESULT 7
ID YK23_YEAST STANDARD; PRT; 271 AA.
AC P36136;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 31.0 KD PROTEIN IN GAP1-NAP1 INTERGENIC REGION.
GN YK043C.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RA URRESTARAZU L.A., JAUNIAUX J.-C.;
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Z28268; G486491; -.
DR PIR; S38115; S38115.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 271 AA; 31022 MW; F8A036A8 CRC32;

Query Match 74.2%; Score 46; DB 11; Length 271;
Best Local Similarity 55.6%; Pred. No. 7.36e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 233 daggigvls 241
QY 1 EAAGIGILT 9

RESULT 8
ID GLMU_ECOLI STANDARD; PRT; 456 AA.
AC P17114;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE UDP-N-ACETYLGLUCOSAMINE PYROPHOSPHORYLASE (EC 2.7.7.23) (N-
DE ACETYLGLUCOSAMINE-1-PHOSPHATE URIDYLTRANSFERASE).
GN GLMU.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85121806.
RA WALKER J.E., GAY N.J., SARASTE M., EBERLE A.N.;
RL BIOCHEM. J. 224:799-815(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 93315143.
RA BURLAND V.D., PLUNKETT G. III, DANIELS D.L., BLATTNER F.R.;
RL GENOMICS 16:551-561(1993).
RN [3]
RP IDENTIFICATION.
RX MEDLINE; 94012475.
RA MENGIN-LECREULX D., VAN HEIJENOORT J.;
```

```

RL J. BACTERIOL. 175:6150-6157(1993).
CC -1- FUNCTION: BIFUNCTIONAL ENZYME RESPONSIBLE FOR THE ACETYLATION OF
CC GLC-N-1-P TO GIVE GLCNAC-1-P AND THE SYNTHESIS OF UDP-GLCNAC.
CC -1- CATALYTIC ACTIVITY: UTP + N-ACETYL-ALPHA-D-GLUCOSAMINE
CC 1-PHOSPHATE = PYROPHOSPHATE + UDP-N-ACETYL-D-GLUCOSAMINE.
CC -1- PATHWAY: PEPTIDOGLYCAN AND LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT THAT CREATES TWO ORFS.
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
DR EMBL: X01631; G43267; -.
DR EMBL: L10328; G290579; ALT_FRAME.
DR EMBL: L10328; G290578; ALT_FRAME.
DR ECGENE; EG11198; GLMU.
DR PROSITE; PS00101; HEXAPEP_TRANSFERASES.
DR PEPTIDOGLYCAN SYNTHESIS; CELL WALL; TRANSFERASE;
KW NUCLEOTIDYLTRANSFERASE; REPEAT; MULTIFUNCTIONAL ENZYME.
SQ SEQUENCE 456 AA; 49162 MW; 7A80D509 CRC32;

Query Match 71.0%; Score 44; DB 4; Length 456;
Best Local Similarity 75.0%; Pred. No. 1.96e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 124 9949lltv 131
:|:|:|:|
QY 3 AGIGILTV 10

RESULT 9
ID FLID_ECOLI STANDARD; PRT; 467 AA.
AC P24216;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE FLAGELLAR HOOK-ASSOCIATED PROTEIN 2 (HAP2) (FILAMENT CAP PROTEIN).
GN FLID OR FLBC OR FLAV.
OS ESCHERICHIA COLI.
CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
CC ENTEROBACTERIACEAE.
[1]
SEQUENCE FROM N.A.
RC STRAIN-JA11;
RX MEDLINE; 92407478.
RA KAWAGISHI I., MUELLER V., WILLIAMS A.W., IRIKURA V.M., MACNAB R.M.;
RL J. GEN. MICROBIOL. 138:1051-1065(1992).
[2]
SEQUENCE OF 1-152 FROM N.A.
RC STRAIN-K12;
RX MEDLINE; 89281489.
RA HANAFUSA T., SAKAI A., TOMINAGA A., ENOMOTO M.;
RL MOL. GEN. GENET. 216:44-50(1989).
[3]
SEQUENCE OF 1-8 FROM N.A.
RX MEDLINE; 83238225.
RA SZEKELY E., SIMON M.;
RL J. BACTERIOL. 155:74-81(1983).
CC -1- FUNCTION: CAPPING PROTEIN FOR THE FLAGELLA; FORMS THE DISTAL END
CC OF THE FLAGELLA.
CC -1- SIMILARITY: TO OTHER FILAMENT CAP PROTEINS.
DR EMBL; M85240; G145988; -.
DR EMBL; X17440; G41650; -.
DR EMBL; J01607; E17475; -.
DR PIR; PV0005; PV0005.
DR ECGENE; EG10841; FLID.
FT INIT_MET 0
FT CONFLICT 113 113 BY SIMILARITY.
SQ SEQUENCE 467 AA; 48270 MW; 14800A2E CRC32;

Query Match 71.0%; Score 44; DB 4; Length 467;
Best Local Similarity 50.0%; Pred. No. 1.96e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 368 dasvgaliv 377

```

```

:|:|:|:|
QY 1 EAAGIGILTV 10

RESULT 10
ID XYND_BACPO STANDARD; PRT; 635 AA.
AC P45796;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE ENDO-1,4-BETA-XYLANASE D PRECURSOR (EC 3.2.1.8) (XYLANASE D)
DE (1,4-BETA-D-XYLAN XYLANOXYDROLASE D).
GN XYND.
OS BACILLUS POLYMYXA.
CC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
[1]
SEQUENCE FROM N.A.
RC STRAIN-ATCC 842;
RX MEDLINE; 92041687.
RA GOSALBES M.J., PEREZ-GONZALEZ J.A., GONZALEZ R., NAVARRO A.;
RL J. BACTERIOL. 173:7705-7710(1991).
CC -1- FUNCTION: SHOWS XYLANASE ACTIVITY AS WELL AS ALPHA-L-
CC ARABINOFURANOSIDASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC LINKAGES IN XYLANS.
CC -1- PATHWAY: XYLAN DEGRADATION.
CC -1- SIMILARITY: BELONGS TO FAMILY 43 OF GLYCOSYL HYDROLASES.
DR EMBL; X57094; G48816; -.
KW XYLAN DEGRADATION; HYDROLASE; GLYCOSIDASE; SIGNAL.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 635 XYLANASE D.
SQ SEQUENCE 635 AA; 67914 MW; 078AAB82 CRC32;

Query Match 71.0%; Score 44; DB 10; Length 635;
Best Local Similarity 75.0%; Pred. No. 1.96e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 149 gagigvlt 156
:|:|:|:|
QY 2 AAGIGILT 9

RESULT 11
ID ATPEL_ARATH STANDARD; PRT; 132 AA.
AC P09468;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE ATP SYNTHASE EPSILON CHAIN (EC 3.6.1.34).
GN ATPE.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OG CHLOROPLAST.
CC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
CC CAPPARALES; CRUCIFERAE.
[1]
SEQUENCE FROM N.A.
RC STRAIN-CV. LANDSBERG ERECTA;
RX MEDLINE; 89057486.
RA CHEN H.-C., WINTZ H., WEIL J.-H., PILLAY D.T.N.;
RL NUCLEIC ACIDS RES. 16:10372-10372(1988).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE.
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.
CC EMBL; X12889; G11334; -.
DR PIR; S01903; S01903.
KW ATP SYNTHESIS; CHLOROPLAST; THYLAKOID MEMBRANE; CF(1);
KW HYDROLASE; HYDROGEN ION TRANSPORT.
SQ SEQUENCE 132 AA; 14472 MW; D826F274 CRC32;

```

```

Query Match          69.4%; Score 43; DB 1; Length 132;
Best Local Similarity 66.7%; Pred. No. 3.15e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 43 avdgilgti 51
   | | | | |
QY 2 AAGIGILTV 10

RESULT 12
ID RL4_BACSU STANDARD; PRT; 207 AA.
AC P42921;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE 50S RIBOSOMAL PROTEIN L4.
GN RPDL.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA YOSHIKAWA H., YASUMOTO K., TAKAHASHI H.;
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=SG38;
RA LI X., LINDAHL L., ZENGEL J.M.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -|- FUNCTION: THIS PROTEIN BINDS DIRECTLY AND SPECIFICALLY TO 23S
CC RNA (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE L4P FAMILY OF RIBOSOMAL PROTEINS.
DR EMBL; D50302; G786154; -.
DR EMBL; U43929; G1165304; -.
DR SUBTILIST; BG11219; RPDL.
KW RIBOSOMAL PROTEIN; RNA-BINDING.
SQ SEQUENCE 207 AA; 22391 MW; 920ED2AF CRC32;

Query Match          69.4%; Score 43; DB 8; Length 207;
Best Local Similarity 60.0%; Pred. No. 3.15e+01;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 177 eanginlvld 186
   | | | | |
QY 1 EAAGIGILTV 10

RESULT 13
ID IAP_ECOLI STANDARD; PRT; 345 AA.
AC P10423;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE ALKALINE PHOSPHATASE ISOZYME CONVERSION PROTEIN PRECURSOR
DE (EC 3.4.11.-).
GN IAP.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88058748.
RA ISHINO Y., SHINAGAWA H., MAKINO K., AMEMURA M., NAKATA A.;
RL J. BACTERIOL. 169:5429-5433(1987).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA PLUNKETT G. III;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
[3]
RP SEQUENCE OF 1-111 FROM N.A.
RC STRAIN-K12;

```

```

RX MEDLINE; 92268080.
RA LEYH T.S., VOGT T.F., SUO Y.;
RL J. BIOL. CHEM. 267:10405-10410(1992).
CC -|- FUNCTION: THIS PROTEIN, PRESUMABLY AN AMINOPEPTIDASE, MEDIATES THE
CC CONVERSION OF E. COLI ALKALINE PHOSPHATASE ISOZYME 1, TO ISOZYMES 2
CC AND 3 BY REMOVING, ONE BY ONE, THE TWO AMINO-TERMINAL ARGININE
CC RESIDUES.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY U2.
DR EMBL; M18270; G146430; -.
DR EMBL; U29579; G882646; -.
DR EMBL; M74586; G551798; -.
DR PIR; A28382; KZEC.
DR ECGENE; EGI0488; IAP.
KW HYDROLASE; AMINOPEPTIDASE; SIGNAL.
FT SIGNAL 1 24 PROBABLE.
FT CHAIN 25 345 ALKALINE PHOSPHATASE ISOZYME CONVERSION
FT PROTEIN.
SQ SEQUENCE 345 AA; 37920 MW; 1FCD6521 CRC32;

```

```

Query Match          69.4%; Score 43; DB 5; Length 345;
Best Local Similarity 50.0%; Pred. No. 3.15e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 267 dkagiavlsv 276
   : | | | | |
QY 1 EAAGIGILTV 10

```

```

RESULT 14
ID YXBC_BACSU STANDARD; PRT; 388 AA.
AC P46333;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN IN HTPG-IOLR INTERGENIC
DE REGION.
GN YXBC OR SS92BR.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE; 96093926.
RA YOSHIDA K.-I., SEKI S., FUJIMURA M., MIWA Y., YOSHIDA K.-I.;
RL DNA RES. 2:61-69(1995).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; D45242; G904202; -.
DR SUBTILIST; BG11360; YXBC.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2.
KW HYPOTHETICAL PROTEIN; TRANSPORT; TRANSMEMBRANE.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 77 97 POTENTIAL.
FT TRANSMEM 105 125 POTENTIAL.
FT TRANSMEM 140 160 POTENTIAL.
FT TRANSMEM 184 184 POTENTIAL.
FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT TRANSMEM 309 329 POTENTIAL.
FT TRANSMEM 342 362 POTENTIAL.
SQ SEQUENCE 388 AA; 42163 MW; 7E9AEDF4 CRC32;

```

```

Query Match          69.4%; Score 43; DB 11; Length 388;
Best Local Similarity 85.7%; Pred. No. 3.15e+01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 285 giginlv 291
   | | | | |
QY 4 GIGILTV 10

```

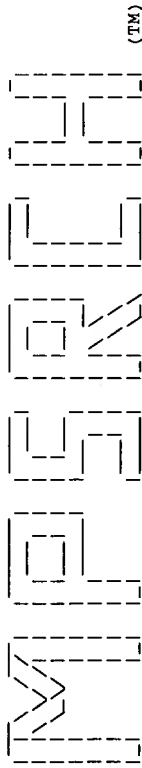
```
RESULT 15
ID ACHE_MOUSE STANDARD; PRT; 493 AA.
AC P20782;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ACETYLCHOLINE RECEPTOR PROTEIN, EPSILON CHAIN PRECURSOR.
GN ACRE.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89214211.
RA BUONANNO A., MUDD J., MERLIE J.P.;
RL J. BIOL. CHEM. 264:7611-7616(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91067487.
RA GARDNER P.D.;
RL NUCLEIC ACIDS RES. 18:6714-6714(1990).
CC -1- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC MEMBRANE.
CC -1- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,
CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE
CC MUSCLE) CHAINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; J04698; G387086; -
DR EMBL; X55718; G53161; -
DR PIR; S13592; ACMSE.
DR PIR; B33358; B33358.
DR PROSITE; PS00236; NEUOTR_ION_CHANNEL.
KW RECEPTOR; POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN; SIGNAL;
KW TRANSMEMBRANE.
FT SIGNAL 1 20
FT CHAIN 21 493 ACETYLCHOLINE RECEPTOR PROTEIN, EPSILON.
FT DOMAIN 21 239 EXTRACELLULAR.
FT TRANSMEM 240 264 POTENTIAL.
FT DOMAIN 265 272 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 273 291 POTENTIAL.
FT DOMAIN 292 306 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 307 328 POTENTIAL.
FT DOMAIN 329 456 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 457 480 POTENTIAL.
FT DOMAIN 481 493 EXTRACELLULAR (POTENTIAL).
FT DISULFID 148 162 BY SIMILARITY.
FT CARBOHYD 86 86 POTENTIAL.
FT CARBOHYD 161 161 PROBABLE.
SQ SEQUENCE 493 AA; 54914 MW; BB5BF2C0 CRC32;

Query Match 69.4%; Score 43; DB 1; Length 493;
Best Local Similarity 75.0%; Pred.No. 3.15e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 93 agvgilrv 100
|:|:|:|
Qy 3 AGIGILTV 10

Search completed: Tue Jun 10 11:18:39 1997
Job time : 9 secs.
```

\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPerch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:20:41 1997; MasPar time 1.96 Seconds  
55.486 Million cell updates/sec

Tabular output not generated.

Title: >US-08-231-565A-18  
Description: (1-10) from US08231565A.pep  
Perfect Score: 63

Sequence: 1 AAGIGILTIV 10

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-gensseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 15.667; Variance 49.649; scale 0.316

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	63	100.0	10	15	MART-1 melanoma anti	1.57e+00
2	63	100.0	118	12	Tumour rejection anti	1.57e+00
3	63	100.0	118	15	MART-1 melanoma anti	1.57e+00
4	56	88.9	9	15	MART-1 melanoma anti	9.25e+00
5	56	88.9	10	15	MART-1 melanoma anti	9.25e+00
6	55	87.3	9	15	MART-1 melanoma anti	1.19e+01
7	52	82.5	9	15	Modified MART-1 melan	2.49e+01
8	52	82.5	9	15	Modified MART-1 melan	2.49e+01
9	52	82.5	9	15	Modified MART-1 melan	2.49e+01
10	52	82.5	479	3	Sequence encoded by t	2.49e+01
11	52	82.5	479	1	Complete sequence of	2.49e+01
12	51	81.0	9	15	Modified MART-1 melan	3.17e+01
13	50	79.4	9	15	Modified MART-1 melan	4.05e+01
14	49	77.8	9	15	Modified MART-1 melan	5.15e+01
15	48	76.2	9	15	Modified MART-1 melan	6.55e+01
16	48	76.2	9	15	Modified MART-1 melan	6.55e+01
17	48	76.2	9	15	Modified MART-1 melan	6.55e+01
18	48	76.2	9	15	Modified MART-1 melan	6.55e+01
19	48	76.2	9	15	Modified MART-1 melan	6.55e+01
20	48	76.2	9	15	Modified MART-1 melan	6.55e+01

21	45	71.4	385	2	R10692	Cephalosporin antibio	1.33e+02
22	43	68.3	9	15	R84790	Modified MART-1 melan	2.13e+02
23	43	68.3	25	8	R39763	Melittin delta 4 (del	2.13e+02
24	43	68.3	548	18	W04212	Luciola lateralis luc	2.13e+02
25	43	68.3	548	6	R33710	Thermostable lucifera	2.13e+02
26	43	68.3	548	3	R14124	Luciferase mutant ile	2.13e+02
27	43	68.3	548	1	R03731	Luciola lateralis rec	2.13e+02
28	43	68.3	548	6	R30803	Thermostable lucifera	2.13e+02
29	43	68.3	548	16	R84178	Luciferase of Hotaria	2.13e+02
30	43	68.3	673	18	W04208	Streptavidin/lucifera	2.13e+02
31	43	68.3	685	18	W04209	Mutant streptavidin/1	2.13e+02
32	42	66.7	231	5	R28440	Cephalosporin C acety	2.68e+02
33	42	66.7	817	9	R47201	DPM2 mannosyl transfe	2.68e+02
34	42	66.7	973	13	R75707	Eph-related PKCek9.	2.68e+02
35	42	66.7	1161	14	R78166	Human beta-2 integrin	2.68e+02
36	41	65.1	9	15	R84765	Modified MART-1 melan	3.37e+02
37	41	65.1	9	15	R84798	Modified MART-1 melan	3.37e+02
38	41	65.1	9	15	R84800	Modified MART-1 melan	3.37e+02
39	41	65.1	9	15	R84789	Modified MART-1 melan	3.37e+02
40	41	65.1	362	2	R03144	Sequence of HLA-B*51 a	3.37e+02
41	41	65.1	548	5	R05788	Luciferase.	3.37e+02
42	41	65.1	548	3	R14120	Luciferase mutant Ser	3.37e+02
43	41	65.1	548	3	R14123	Luciferase mutant Asn	3.37e+02
44	41	65.1	1464	18	R80970	Human excitatory amin	3.37e+02
45	41	65.1	1464	10	R35529	Human NMDA R2A recept	3.37e+02

ALIGNMENTS

RESULT 1  
ID R84198 standard; Peptide; 10 AA.  
AC R84198;  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen immunogenic peptide M10-4.  
KW MART-1; M10-4; melanoma antigen recognised by T-cells; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN W09589193-A2.  
PD 02-NOV-1995.  
PF 11-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 12; Page 122; 184pp; English.  
CC Immunogenic peptide M10-4 is a derivative of peptide M9-2 (R84196)  
CC which is based on the melanoma antigen (MART-1) (see R84212).  
CC M9-2 may be modified to improve immunogenicity (see R84783-R84800)  
CC and used in medicaments for the treatment or prevention (by  
CC immunization) of melanoma. Antibodies against MART-1 and its  
CC immunogenic peptides may be used in the detection and isolation of  
CC MART-1 from a sample, the detection of which is indicative of a  
CC disease state (melanoma or metastatic melanoma).  
SQ See also R84196.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 63; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.57e+00;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aagigiltvi 10

Qy 1 AAGIGILTIV 10

RESULT 2

ID R63158 standard; Protein; 118 AA.

AC R63158:  
DE Tumour rejection antigen precursor.  
KW Tumour rejection antigen; precursor; HLA-A2 molecule; tyrosinase;  
KW isolation; melanoma; cell line; LB-39-MEL; diagnosis; vaccine;  
KW therapy.  
OS Homo sapiens.  
PN W0942126-A.  
PD 23-SEP-1994.  
PF 09-MAR-1994; U02487.  
PR 18-MAR-1993; US-032978.  
PA (LUDW-) LUDWIG INST CANCER RES.  
PI Boon-Falleur T, Brichard V, De Plaen E, Traversari C;  
PI Van Pel A, Wolfel T;  
DR WPI: 94-316544/39.  
DR N-PSDB: Q76370.  
PT Nucleic acid coding for a tumour rejection antigen precursor - is  
PT used for developing prods. for diagnosis or treatment of expression  
PT related disorders, partic. melanoma  
PS Claim 5; Page 14; 26pp; English.  
CC This sequence represents the tumour rejection antigen precursor which is  
CC processed to a tumour rejection antigen presented by HLA-A2 molecules.  
CC The tumour rejection antigen is not related to tyrosinase. The cDNA  
CC encoding this sequence was isolated from the melanoma cell line,  
CC LB-39-MEL. The tumour rejection antigen may be used for diagnosis or  
CC in vaccines or for therapy of disorders characterised by the expression  
CC of the tumour rejection antigen precursor, particularly melanoma.  
SQ Sequence 118 AA;

Query Match 100.0%; Score 63; DB 12; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.57e+00;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 aagigiltvi 36  
|||||  
QY 1 AAGIGILTVI 10

RESULT 3  
ID R84212 standard; Protein; 118 AA.  
AC R84212;  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen.  
KW MART-1; melanoma antigen recognised by T-cell; melanoma;  
KW metastatic melanoma; tumour-associated antigen; immunogen;  
KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
OS Mammalian.  
FT Key  
FT Region 27..47  
FT Note: "hydrophobic region"  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
PI WPI: 95-382963/49.  
DR N-PSDB: T02714.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 11; Page 117; 184pp; English.  
CC The melanoma antigen (MART-1) is produced by recombinant DNA  
CC methods, i.e. preferably using a baculovirus vector for expression  
CC in insect cell cultures. MART-1 protein is a source of immunogenic  
CC peptides (see R84196 for peptide M9-2) which are optionally modified  
CC (see R84783-R84800) and used in medicaments for the treatment or  
CC prevention (by immunization) of melanoma. Antibodies against MART-1  
CC and its immunogenic peptides may be used in the detection and  
CC isolation of MART-1 from a sample, the detection of which is  
CC indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 118 AA;

Query Match 100.0%; Score 63; DB 15; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.57e+00;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 aagigiltvi 36  
|||||  
QY 1 AAGIGILTVI 10

RESULT 4  
ID R84196 standard; Peptide; 9 AA.  
AC R84196;  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen immunogenic peptide M9-2.  
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
PI WPI: 95-382963/49.  
DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
DR vectors, host cells and antibodies, used to detect, treat and  
DR immunise animal against melanoma.  
PS Claim 12; Page 117; 184pp; English.  
CC Immunogenic peptide M9-2 is based on the melanoma antigen (MART-1)  
CC (see R84212). M9-2 may be modified to improve immunogenicity  
CC (see R84783-R84800) and used in medicaments for the treatment or  
CC prevention (by immunization) of melanoma. Antibodies against MART-1  
CC and its immunogenic peptides may be used in the detection and  
CC isolation of MART-1 from a sample, the detection of which is  
CC indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 88.9%; Score 56; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.25e+00;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aagigiltv 9  
|||||  
QY 1 AAGIGILTV 9

RESULT 5  
ID R84197 standard; Peptide; 10 AA.  
AC R84197;  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen immunogenic peptide M10-3 derivative.  
KW MART-1; M10-3; melanoma antigen recognised by T-cells; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
PI WPI: 95-382963/49.  
DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
DR vectors, host cells and antibodies, used to detect, treat and  
DR immunise animal against melanoma.  
PS Claim 12; Page 122; 184pp; English.  
CC Immunogenic peptide M10-3 is a derivative of peptide M9-2 (R84196)

CC which is based on the melanoma antigen (MART-1) (see R84212).  
 CC M9-2 may be modified to improve immunogenicity (see R84783-R84800).  
 CC and used in medicaments for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against MART-1 and its  
 CC immunogenic peptides may be used in the detection and isolation of  
 CC MART-1 from a sample, the detection of which is indicative of a  
 CC disease state (melanoma or metastatic melanoma).  
 CC See also R84198  
 CC Sequence 10 AA;

Query Match 88.9%; Score 56; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 9.25e+00;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 aagigiltv 10  
 |||||  
 Qy 1 AAGIGILTV 9

RESULT 6  
 ID R84764 standard; Peptide; 9 AA.  
 AC R84764;  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen immunogenic peptide M9-3.  
 KW MART-1; M9-3; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09529193-A2.  
 PD 04-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 2; Page 72; 184pp; English.  
 CC Immunogenic peptide M9-3 is based on the melanoma antigen (MART-1)  
 CC (see R84212). M9-3 may be modified to improve immunogenicity  
 CC (see R84763-R84782) and used in medicaments (vaccines) for  
 CC treatment or prevention (by immunization) of melanoma.  
 CC Antibodies against MART-1 and its immunogenic peptides may be used  
 CC in the detection and isolation of MART-1 from a sample, the  
 CC detection of which is indicative of a disease state (melanoma  
 CC or metastatic melanoma).  
 CC See also R84196-R84198.  
 CC Sequence 9 AA;

Query Match 87.3%; Score 55; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.19e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 gigtltvi 8  
 |||||  
 Qy 3 GIGILTVI 10

RESULT 7  
 ID R84787 standard; Peptide; 9 AA.  
 AC R84787;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1F.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09529193-A2.  
 PD 02-NOV-1995.

PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 CC Sequence 9 AA;

Query Match 82.5%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.49e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||  
 Qy 2 AGIGILTV 9

RESULT 8  
 ID R84788 standard; Peptide; 9 AA.  
 AC R84788;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1Y.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 CC Sequence 9 AA;

Query Match 82.5%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.49e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||  
 Qy 2 AGIGILTV 9

RESULT 9  
 ID R84786 standard; Peptide; 9 AA.  
 AC R84786;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1W.



KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; US-231565.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 82.5%; Score 52; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.49e+01;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 agigiltv 9  
 |||||  
 QY 2 AGIGILTV 9

RESULT 10  
 ID P50034 standard; Protein; 479 AA.  
 AC P50034;  
 DT 07-SEP-1991 (first entry)  
 DE Sequence encoded by the coding region of the PRV gIII gene.  
 KW Sub-unit vaccine; antigen; immunoreactive determinant.  
 OS Pseudorabies virus (PRV).  
 PN EP-162738-A.  
 PD 27-OCT-1985.  
 PF 09-APR-1985; 400704.  
 PR 09-APR-1984; US-598073.  
 PR 03-APR-1985; US-719773.  
 PA (MOLE-) MOLECULAR GENETICS.  
 PI Robbins AK, Watson RJ, Enquist LW;  
 DR WPI; 85-298083/48.  
 DR N-PSDB; N50035.  
 PT Prodn. of Pseudorabies virus sub-unit vaccines - useful for  
 PT conferring protection against the virus infections and for  
 PT diagnosis esp. with sheep sera  
 PS Example; Fig 3; 153pp; English.  
 CC The inventors claim a recombinant vector comprising a DNA sequence  
 CC coding for an immunoreactive and antigenic determinant of a  
 CC pseudorabies virus protein, and a polypeptide having an  
 CC immunoreactive and antigenic determinant of a Pseudorabies virus  
 CC protein. Vaccines contg. the polypeptide are obtd. economically and  
 CC in large amts., for use in conferring protection against  
 CC Pseudorabies virus.  
 SQ Sequence 479 AA;

Query Match 82.5%; Score 52; DB 3; Length 479;  
 Best Local Similarity 66.7%; Pred. No. 2.49e+01;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigilaiv 464  
 |||||  
 QY 2 AGIGILTVI 10

RESULT 11  
 ID P81013 standard; protein; 479 AA.  
 AC P81013;  
 DT 14-DEC-1990 (first entry)  
 DE Complete sequence of the pseudorabies virus (PRV) gIII glycoprotein  
 DE gene (noncoding strand)  
 KW Pseudorabies virus; live viral vaccine; antigen; diagnosis;  
 KW nonessential region.  
 OS Pseudorabies virus.  
 PN EP-256677-A.  
 PD 24-FEB-1988.  
 PF 17-JUL-1987; 306364.  
 PR 18-JUL-1986; US-886691.  
 PA (DUPO) Du Pont De Nemours Co.  
 PI Enquist LW, Robbins AK, Whealy ME;  
 DR WPI; 88-051443/08.  
 DR N-PSDB; N81334.  
 PT Recombinant pseudo-rabies -  
 PT having non-essential genome regions deleted and pot. having  
 PT foreign DNA fragments inserted  
 PS Disclosure; p; English.  
 CC XhoI/gIII region and any other nonessential region of the PV genome can  
 CC be manipulated to produce the deletions and insertions by standard  
 CC recombinant DNA techniques. Recombinants can be used for the prodn. of  
 CC desired proteins. Antigens produced are useful in diagnostic kits  
 CC and as vaccines (claimed). also, the PV engineered to contain the genes  
 CC for appropriate antigens also should have utility as a live vaccine.  
 SQ Sequence 479 AA;

Query Match 82.5%; Score 52; DB 1; Length 479;  
 Best Local Similarity 66.7%; Pred. No. 2.49e+01;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigilaiv 464  
 |||||  
 QY 2 AGIGILTVI 10

RESULT 12  
 ID R84785 standard; Peptide; 9 AA.  
 AC R84785;  
 DT 25-APR-1996 (first entry)  
 DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2I.  
 KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Example 5; Page 105; 184pp; English.  
 CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is  
 CC an immunogenic peptide based on the melanoma antigen (MART-1) (see  
 CC R84212). The peptides are used in medicaments for the treatment or  
 CC prevention (by immunization) of melanoma. Antibodies against MART-1  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of MART-1 from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 9 AA;

Query Match 81.0%; Score 51; DB 15; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 3.17e+01;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Db 1 aigigiltv 9
QY 1 AAGIGILTV 9

RESULT 13
ID R84784 standard; Peptide; 9 AA.
AC R84784;
DE 25-APR-1996 (first entry)
DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2M.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 105; 184pp; English.
CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is
CC an immunogenic peptide based on the melanoma antigen (MART-1) (see
CC R84212). The peptides are used in medicaments for the treatment or
CC prevention (by immunization) of melanoma. Antibodies against MART-1
CC and its immunogenic peptides may be used in the detection and
CC isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 79.4%; Score 50; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.05e+01;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 aigigiltv 9
QY 1 AAGIGILTV 9

RESULT 14
ID R84783 standard; Peptide; 9 AA.
AC R84783;
DE 25-APR-1996 (first entry)
DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-2L.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 105; 184pp; English.
CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is
CC an immunogenic peptide based on the melanoma antigen (MART-1) (see
CC R84212). The peptides are used in medicaments for the treatment or
CC prevention (by immunization) of melanoma. Antibodies against MART-1
CC and its immunogenic peptides may be used in the detection and
CC isolation of MART-1 from a sample, the detection of which is

CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 76.2%; Score 48; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.55e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigitlv 9
QY 3 GIGILTV 9

CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 77.8%; Score 49; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 5.15e+01;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 aigigiltv 9
QY 1 AAGIGILTV 9

RESULT 15
ID R84795 standard; Peptide; 9 AA.
AC R84795;
DE 25-APR-1996 (first entry)
DE Modified MART-1 melanoma antigen immunogenic peptide M9-2-1W2L.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 105; 184pp; English.
CC R84783-800 are M9-2 peptides modified to improve immunogenicity. M9-2 is
CC an immunogenic peptide based on the melanoma antigen (MART-1) (see
CC R84212). The peptides are used in medicaments for the treatment or
CC prevention (by immunization) of melanoma. Antibodies against MART-1
CC and its immunogenic peptides may be used in the detection and
CC isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 76.2%; Score 48; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.55e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 gigitlv 9
QY 3 GIGILTV 9

Search completed: Tue Jun 10 11:20:50 1997
Job time : 9 secs.

```

\*\*\*\*\*  
W O R L D  
\*\*\*\*\*  
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:20:12 1997; MasPar time 2.78 Seconds  
102.637 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-231-565A-18  
Description: (1-10) from US08231565A.pep  
Perfect Score: 63  
Sequence: 1 AAGIGILTVI 10

Scoring table: PAM 150  
Gap 15  
Searched: 89912 seqs, 28507787 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir50  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unann16:unrev  
Statistics: Mean 22.523; Variance 32.002; scale 0.704

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description	Pred. No.
1	63	100.0	118 13	A55253	118 13		melanoma antigen MAR	2.43e-02
2	63	100.0	118 13	I38506	118 13		melan-A protein - hu	2.43e-02
3	52	82.5	479 4	VGBEPB	479 4		glycoprotein gIII pr	3.58e+00
4	48	76.2	231 12	S48276	231 12		YSA1 protein - yeast	1.91e+01
5	47	74.6	101 10	A33551	101 10		H+-transporting ATP	2.87e+01
6	47	74.6	308 9	H64153	308 9		hypothetical protein	2.87e+01
7	46	73.0	637 11	S35221	637 11		globulin Bg1l precu	4.28e+01
8	45	71.4	339 16	S62369	339 16		methylobalammin: Coe	6.35e+01
9	45	71.4	385 16	S54103	385 16		beta-lactamase - Lys	6.35e+01
10	45	71.4	420 14	S59131	420 14		kan-1 protein - rat	6.35e+01
11	45	71.4	552 9	S25200	552 9		narQ protein - Esche	6.35e+01
12	45	71.4	566 9	A46162	566 9		nitrate sensor prote	6.35e+01
13	44	69.8	394 10	A55045	394 10		hydroxylacyl-CoA deh	9.35e+01
14	44	69.8	611 8	JT0592	611 8		hypothetical protein	9.35e+01
15	44	69.8	635 10	S19011	635 10		endo-1,4-beta-xylana	9.35e+01
16	44	69.8	1091 3	IJCHNL	1091 3		neural cell adhesion	9.35e+01
17	43	68.3	19 10	A34624	19 10		31k flagellin - Meth	1.37e+02
18	43	68.3	103 14	S35077	103 14		cellulobrevin - rat	1.37e+02
19	43	68.3	109 8	I39724	109 8		ORF13a - Agrobacteri	1.37e+02
20	43	68.3	132 5	S01903	132 5		H+-transporting ATP	1.37e+02
21	43	68.3	199 4	WM5EX4	199 4		UL4 protein - human	1.37e+02

22	43	68.3	201 4	WMBEHL	UL4 protein - human	1.37e+02
23	43	68.3	216 10	C41316	flagellin B2 precurs	1.37e+02
24	43	68.3	218 10	B41316	flagellin B1 precurs	1.37e+02
25	43	68.3	220 12	S40931	hypothetical protein	1.37e+02
26	43	68.3	394 8	C64185	cell division protein	1.37e+02
27	43	68.3	401 16	S62465	hypothetical protein	1.37e+02
28	43	68.3	493 3	ACMSE	nicotinic acetylcholi	1.37e+02
29	43	68.3	536 8	A55544	flagellar basal body	1.37e+02
30	43	68.3	548 6	S33788	luciferase - southern	1.37e+02
31	43	68.3	548 2	S23437	Photinus-luciferin 4	1.37e+02
32	43	68.3	548 12	S57417	luciferase - Japanes	1.37e+02
33	43	68.3	664 11	S53037	PIB1 protein - yeast	1.37e+02
34	43	68.3	1108 14	A55915	guanylate cyclase (E	1.37e+02
35	43	68.3	1110 14	S55279	guanylate cyclase pr	1.37e+02
36	43	68.3	1143 9	I84547	mdl protein - Escher	1.37e+02
37	42	66.7	110 4	VHBPDL	major capsid protein	1.99e+02
38	42	66.7	180 10	H64342	hypothetical protein	1.99e+02
39	42	66.7	286 13	S08993	signal sequence rece	1.99e+02
40	42	66.7	300 10	S08244	hypothetical protein	1.99e+02
41	42	66.7	361 10	S47170	hypothetical protein	1.99e+02
42	42	66.7	461 9	G64107	thiophene and furan	1.99e+02
43	42	66.7	477 10	H64210	eggshell protein p48	1.99e+02
44	42	66.7	501 10	S18573	L-lysine transport p	1.99e+02
45	42	66.7	521 12	A53153	glucose transport pr	1.99e+02

ALIGNMENTS

RESULT 1  
ENTRY A55253 #type complete  
TITLE melanoma antigen MART-1 - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change  
ACCESSIONS A55253  
REFERENCE A55253  
#authors Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.;  
Rivoltini, L.; Topalian, S.L.; Miki, T.; Rosenberg, S.A.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1994) 91:3515-3519  
#title Cloning of the gene coding for a shared human melanoma  
antigen recognized by autologous T cells infiltrating into  
tumor.  
#accession A55253  
#status preliminary  
#molecule\_type mRNA  
#residues 1-118 #label KAW  
#cross-references GB:006452  
SUMMARY #length 118 #molecular-weight 13157 #checksum 3535

Query Match 100.0%; Score 63; DB 13; Length 118;  
Best Local Similarity 100.0%; Pred. No. 2.43e-02;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	27	aagigiltvi 36	
Qy	1	AAGIGILTVI 10	
RESULT 2			
ENTRY I38506 #type complete			
TITLE melan-A protein - human			
ORGANISM #formal_name Homo sapiens #common_name man			
DATE 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change			
ACCESSIONS I38506			
REFERENCE I38506			
#authors Coullie, P.G.; Brichard, V.; Van Pel, A.; Wolfel, T.; Schneider, J.; Traversari, C.; Mattei, S.; De Plaen, E.; Lurquin, C.; Szikora, J.P.; Renauld, J.; Boon, T. J. Exp. Med. (1994) 180:35-42 A new gene coding for a differentiation antigen recognized by autologous cytolytic T lymphocytes on HLA-A2 melanomas [see comments].			

```

#cross-references MUID:94275389
#accession I38506
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-118 #label RES
#cross-references EMBL:U06654; NID:g517022; CDS_PID:g517023
SUMMARY #length 118 #molecular-weight 13157 #checksum 3535

Query Match 100.0%; Score 63; DB 13; Length 118;
Best Local Similarity 100.0%; Pred. No. 2.43e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 aagigiltvi 36
|||||
QY 1 AAGIGILTVI 10

RESULT 3
ENTRY VGBEPB #type complete
TITLE glycoprotein gIII precursor - suid herpesvirus 1
ORGANISM #formal_name suid herpesvirus 1
DATE 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change
05-Jan-1996
ACCESSIONS A26097
REFERENCE Robbins, A.K.; Watson, R.J.; Whealy, M.E.; Hays, W.W.;
Enquist, L.W.
#journal J. Virol. (1986) 58:339-347
#title Characterization of a Pseudorabies virus glycoprotein gene
with homology to herpes simplex virus type 1 and type 2
glycoprotein C.
#cross-references MUID:86200375
#accession A26097
#molecule_type DNA
#residues 1-479 #label ROB
#experimental_source strain Becker
CLASSIFICATION #superfamily herpesvirus glycoprotein F
KEYWORDS glycoprotein
FEATURE
1-22 #domain signal sequence #status predicted #label SIG\
23-479 #product glycoprotein gIII #status predicted #label GPG\
40,84,169,192,220, #binding_site carbohydrate (Asn) (covalent) #status
228,285,302 predicted
SUMMARY #length 479 #molecular-weight 51206 #checksum 1630

Query Match 82.5%; Score 52; DB 4; Length 479;
Best Local Similarity 66.7%; Pred. No. 3.58e+00;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigilaiv 464
|||||
QY 2 AGIGILTVI 10

RESULT 4
ENTRY S48276 #type complete
TITLE YSA1 protein - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein YBR0907; protein YBR11c
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995
ACCESSIONS S48276; S45979; S25364; S44691
REFERENCE S48255
#authors Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann,
H.
#journal Yeast (1994) 10:1363-1381
#title Analysis of a 70 kb region on the right arm of yeast
chromosome II.
#accession S48276
#molecule_type DNA
#residues 1-231 #label MAN
#cross-references EMBL:X78993

#note nucleotide sequence is not given
S45927
REFERENCE Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
#authors submission
#accession S45979
#molecule_type DNA
#residues 1-231 #label FE2
#cross-references EMBL:Z35980
REFERENCE S25364
#authors Mannhaupt, G.; Stucka, R.; Ehnlé, S.; Vetter, I.; Feldmann,
H.
#journal Yeast (1992) 8:397-408
#title Molecular analysis of yeast chromosome II between CMD1 and
LYS2: the excision repair gene RAD16 located in this region
belongs to a novel group of double-finger proteins.
#accession S25364
#molecule_type DNA
#residues 1-47 #label MAW
#cross-references EMBL:X66247
GENETICS LISTA:YSA1
#gene
#map_position 2R
SUMMARY #length 231 #molecular-weight 26087 #checksum 4809

Query Match 76.2%; Score 48; DB 12; Length 231;
Best Local Similarity 75.0%; Pred. No. 1.91e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 79 gigitlil 86
|||||
QY 3 GIGILTVI 10

RESULT 5
ENTRY A33351 #type complete
TITLE H+-transporting ATP synthase (EC 3.6.1.34) proteolipid chain
- Sulfolobus acidocaldarius
ORGANISM #formal_name Sulfolobus acidocaldarius
DATE 20-Dec-1989 #sequence_revision 20-Dec-1989 #text_change
23-Jun-1993
ACCESSIONS A33351
REFERENCE Denda, K.; Konishi, J.; Oshima, T.; Date, T.; Yoshida, M.
#authors J. Biol. Chem. (1989) 264:7119-7121
#title A gene encoding the proteolipid subunit of Sulfolobus
acidocaldarius ATPase complex.
#cross-references MUID:89214142
#accession A33351
#status preliminary
#molecule_type DNA
#residues 1-101 #label DEN
#cross-references GB:J04740
KEYWORDS hydrolase
SUMMARY #length 101 #molecular-weight 10362 #checksum 4300

Query Match 74.6%; Score 47; DB 10; Length 101;
Best Local Similarity 87.5%; Pred. No. 2.87e+01;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 59 aagigvlt 66
|||||
QY 1 AAGIGILT 8

RESULT 6
ENTRY H64153 #type complete
TITLE hypothetical protein HI0509 - Haemophilus influenzae (strain
Rd KW20)
ORGANISM #formal_name Haemophilus influenzae
DATE 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
10-May-1996
ACCESSIONS H64153
REFERENCE A64000

```

```

#authors Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;
Kirkness, E.F.; Kerlavage, A.R.; Bult, C.J.; Tomb, J.F.;
Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.;
Fitzhugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton,
M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saudak,
D.W.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann,
J.L.; Geoghagen, N.S.M.; Gnehm, C.L.; McDonald, L.A.;
Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
#journal Science (1995) 269:496-512
#title Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.
#accession H64153
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-308 #label TIGR
#cross-references GB:L42023; TIGR:HI0509
#note start codon was translated as Val
#note best homolog was a hypothetical protein from Escherichia
coli
SUMMARY #length 308 #molecular-weight 33345 #checksum 9134
Query Match 74.6%; Score 47; DB 9; Length 308;
Best Local Similarity 55.6%; Pred. No. 2.87e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 132 aglqllaiv 140
||:||||:
QY 2 AGIGILTVI 10
RESULT 7
ENTRY S35221 #type complete
TITLE globulin Bgl precursor - barley
ORGANISM #formal_name Hordeum vulgare #common_name barley
DATE 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
03-Feb-1994
ACCESSIONS S35221
REFERENCE S35221
#authors Heck, G.R.; Chamberlain, A.K.; Ho, T.H.D.
#journal Mol. Gen. Genet. (1993) 239:209-218
#title Barley embryo globulin 1 gene, Bgl1: Characterization of
cDNA, chromosome mapping and regulation of expression.
#accession S35221
#molecule_type mRNA
#residues 1-637 #label HEC
#cross-references EMBL:M64372
GENETICS
#gene Bgl1
#map_position 4
KEYWORDS glycoprotein
FEATURE
174-190
#product globulin Bgl1 #status predicted #label MAT
SUMMARY #length 637 #molecular-weight 72252 #checksum 3309
Query Match 73.0%; Score 46; DB 11; Length 637;
Best Local Similarity 70.0%; Pred. No. 4.28e+01;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 247 adgeglvtvi 256
| | | | |
QY 1 AAGIGILTVI 10
RESULT 8
ENTRY S62369 #type complete
TITLE methylcobalamin: Coenzyme M methyltransferase (isozyme II)
- Methanosarcina barkeri
ORGANISM #formal_name Methanosarcina barkeri
DATE 20-Jul-1996 #sequence_revision 20-Jul-1996 #text_change
20-Jul-1996

```

```

ACCESSIONS S62369
REFERENCE S62368
#authors Harms, U.; Thauer, R.K.
#journal Eur. J. Biochem. (1996) 235:653-659
#title Methylcobalamin:coenzyme M methyltransferase Isoenzymes MtaA
and MtaB from Methanosarcina barkeri. Cloning, sequencing
and differential transcription of the encoding genes, and
functional overexpression of the mtaA gene in Escherichia
coli.
#accession S62369
#status preliminary
#residues 1-339 #label HAR
#cross-references EMBL:X91894
SUMMARY #length 339 #molecular-weight 36761 #checksum 6431
Query Match 71.4%; Score 45; DB 16; Length 339;
Best Local Similarity 75.0%; Pred. No. 6.35e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 307 agvglltv 314
||:||||
QY 2 AGIGILTV 9
RESULT 9
ENTRY S54103 #type complete
TITLE beta-lactamase - Lysobacter lactamgenus (strain YK90)
ORGANISM #formal_name Lysobacter lactamgenus
DATE 08-Jul-1995 #sequence_revision 08-Jul-1995 #text_change
08-Jul-1995
ACCESSIONS S54103
REFERENCE S54099
#authors Kimura, H.; Izawa, M.; Miyashita, H.; Shimizu, Y.; Sumino,
Y.; Suzuki, M.
#submission submitted to the EMBL Data Library, October 1990
#description Gene cluster involved in the cephalosporin biosynthesis from
Lysobacter lactamgenus YK90.
#accession S54103
#status preliminary
#residues 1-385 #label KIM
#cross-references EMBL:X56660
SUMMARY #length 385 #molecular-weight 41878 #checksum 3376
Query Match 71.4%; Score 45; DB 16; Length 385;
Best Local Similarity 66.7%; Pred. No. 6.35e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 173 pgigmlgvi 181
||:||||
QY 2 AGIGILTVI 10
RESULT 10
ENTRY S59131 #type complete
TITLE Kan-1 protein - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
01-Mar-1996
ACCESSIONS S59131
REFERENCE S59131
#authors Furutani, M.; Arai, S.; Higashitsuji, H.; Mise, M.; Fukumoto,
M.; Takano, S.; Nakayama, H.; Imamura, M.; Fujita, J.
#journal Biochem. J. (1995) 311:203-208
#title Reduced expression of kan-1 (encoding putative bile
acid-CoA-amino acid N-acyltransferase) mRNA in livers of
rats after partial hepatectomy and during sepsis.
#accession S59131
#status preliminary
#residues 1-420 #label FUR
#cross-references EMBL:D43964
SUMMARY #length 420 #molecular-weight 46496 #checksum 4868

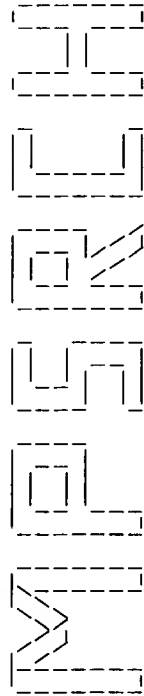
```



#authors Gosalbes, M.J.; Perez-Gonzalez, J.A.; Gonzalez, R.; Navarro,  
A.  
#journal J. Bacteriol. (1991) 173:7705-7710  
#title Two beta-glycanase genes are clustered in Bacillus polymyxa:  
molecular cloning, expression, and sequence analysis of  
genes encoding a xylanase and an endo-beta-(1,3)-(1,  
4)-glucanase.  
#cross-references MUID:92041687  
#accession S19011  
#status preliminary  
#molecule\_type DNA  
#residues 1-635 #label GOS  
#cross-references EMBL:X57094  
#note the authors translated the codon GAA for residue 78 as  
Gly, CCT for residue 272 as Thr, ATC for residue 412  
as Gln, and ATC for residue 478 as Tyr  
glycosidase; hydrolase  
#length 635 #molecular-weight 67914 #checksum 2077  
KEYWORDS  
SUMMARY  
Query Match 69.8%; Score 44; DB 10; Length 635;  
Best Local Similarity 75.0%; Pred. No. 9.35e+01;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 149 gagigvlt 156  
:||||:  
QY 1 AAGIGILT 8

Search completed: Tue Jun 10 11:20:22 1997  
Job time : 10 secs.

\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:19:46 1997; MasPar time 1.97 Seconds  
Tabular output not generated. 107.415 Million cell updates/sec

Title: >US-08-231-565A-18  
Description: (1-10) from US08231565A.pep  
Perfect Score: 63  
Sequence: 1 AAGIGILTIV 10

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 23.598; Variance 27.449; scale 0.860

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	52	82.5	479	10	VGLC_PRIVIF	9.21e+01
2	50	79.4	388	11	YXBC_BACSU	2.48e+00
3	48	76.2	231	11	YSAI_YEAST	6.49e+00
4	47	74.6	101	1	ATPL_SULAC	1.04e+01
5	47	74.6	308	6	MENA_HAEIN	1.04e+01
6	46	73.0	977	11	YD88_SCHPO	1.65e+01
7	45	71.4	566	6	NARQ_ECOLI	2.59e+01
8	44	69.8	278	11	YQHA_BACSU	4.05e+01
9	44	69.8	456	4	GLMB_ECOLI	4.05e+01
10	44	69.8	610	4	FIMB_DICDI	4.05e+01
11	44	69.8	611	11	YD3M_HERAU	4.05e+01
12	44	69.8	635	10	YXND_BACPO	4.05e+01
13	44	69.8	1091	6	NCAL_CHICK	4.05e+01
14	43	68.3	130	11	YHAH_ECOLI	6.27e+01
15	43	68.3	132	1	ATPE_ARATH	6.27e+01
16	43	68.3	199	10	UL04_HSV11	6.27e+01
17	43	68.3	201	10	UL04_HSV2H	6.27e+01
18	43	68.3	216	4	FLA2_METVO	6.27e+01
19	43	68.3	218	4	FLA1_METVO	6.27e+01
20	43	68.3	220	11	Y069_CAEEL	6.27e+01
21	43	68.3	221	9	TRPF_HALVO	6.27e+01
22	43	68.3	394	4	FTSW_HAEIN	6.27e+01

23	43	68.3	401	10	YABA_SCHPO	HYPOTHETICAL 44.4 KD	6.27e+01
24	43	68.3	460	11	YICJ_ECOLI	HYPOTHETICAL 51.0 KD	6.27e+01
25	43	68.3	493	1	ACHE_MOUSE	ACETYLCHOLINE RECEPTO	6.27e+01
26	43	68.3	536	4	FLIF_CAUOR	FLAGELLAR M-RING PROT	6.27e+01
27	43	68.3	548	6	LUCI_LUCLA	LUCIFERIN 4-MONOOXYGE	6.27e+01
28	43	68.3	664	7	PLB1_YEAST	LYSOPHOSPHOLIPASE PRE	6.27e+01
29	43	68.3	1108	3	CYGE_MOUSE	GUANYLYL CYCLASE GC-E	6.27e+01
30	43	68.3	1108	3	CYGE_RAT	GUANYLYL CYCLASE GC-E	6.27e+01
31	43	68.3	1110	3	CYGD_BOVIN	RETINAL GUANYLYL CYCL	6.27e+01
32	43	68.3	1143	6	MDL_ECOLI	MULTIDRUG RESISTANCE-	6.27e+01
33	42	66.7	110	10	VCAD_LAMBD	HEAD DECORATION PROTE	9.61e+01
34	42	66.7	207	8	RL4_BACSU	50S RIBOSOMAL PROTEIN	9.61e+01
35	42	66.7	286	9	SSRA_CANFA	TRANSLOCON-ASSOCIATED	9.61e+01
36	42	66.7	286	9	SSRA_RABIT	TRANSLOCON-ASSOCIATED	9.61e+01
37	42	66.7	300	11	YR33_THEPE	HYPOTHETICAL 33.4 KD	9.61e+01
38	42	66.7	312	2	COX3_BACFI	PROBABLE CYTOCHROME C	9.61e+01
39	42	66.7	332	1	ACOA_ALCEU	ACETABLE:2,6-DICHLOROP	9.61e+01
40	42	66.7	355	11	YCIQ_ECOLI	HYPOTHETICAL 39.3 KD	9.61e+01
41	42	66.7	429	1	AG45_MYLE	46 KD MEMBRANE PROTEI	9.61e+01
42	42	66.7	461	9	THDF_HAEIN	POSSIBLE THIOPHENE AN	9.61e+01
43	42	66.7	530	1	AIP2_YEAST	ACTIN INTERACTING PRO	9.61e+01
44	42	66.7	1530	1	BFRI_SCHPO	BREFELDIN A RESISTANC	9.61e+01
45	42	66.7	1858	7	P3K2_DICDI	PHOSPHATIDYLINOSITOL	9.61e+01

ALIGNMENTS

RESULT 1  
ID VGLC\_PRIVIF STANDARD; PRT; 479 AA.  
AC P06024;  
DT 13-AUG-1987 (REL. 05, CREATED)  
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)  
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
DE GLYCOPROTEIN GIII PRECURSOR.  
OS PSEUDORABIES VIRUS (STRAIN INDIANA-FUNKHAUSER / BECKER) (PRV).  
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 86200375.  
RA ROBBINS A.K., WATSON R.J., WHEALY M.E., HAYS W.W., ENQUIST L.W.;  
J. VIROL. 58:339-347(1986).  
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN C.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN GENE SUPERFAMILY.  
DR EMBL; M12778; G334050; -.  
KW GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.  
FT SIGNAL 1 22  
FT CHAIN 23 479 GLYCOPROTEIN GIII.  
FT CARBOHYD 40 40 POTENTIAL.  
FT CARBOHYD 84 84 POTENTIAL.  
FT CARBOHYD 169 169 POTENTIAL.  
FT CARBOHYD 192 192 POTENTIAL.  
FT CARBOHYD 220 220 POTENTIAL.  
FT CARBOHYD 228 228 POTENTIAL.  
FT CARBOHYD 285 285 POTENTIAL.  
FT CARBOHYD 302 302 POTENTIAL.  
SQ SEQUENCE 479 AA; 51206 MW; 42EE5703 CRC32;

Query Match 82.5%; Score 52; DB 10; Length 479;  
Best Local Similarity 66.7%; Pred. No. 9.21e+01;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 456 agigilaiv 464  
QY 2 AGIGILTIV 10

RESULT 2	STANDARD; PRT; 388 AA.
ID YXBC_BACSU	
AC P46333;	
DT 01-NOV-1995 (REL. 32, CREATED)	
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)	
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)	



DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN IN HTPG-IOLR INTERGENIC  
DE REGION.  
GN YXBC OR SS92BR.  
OC BACILLUS SUBTILIS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168 / BGSCIAL;  
RX MEDLINE; 96093926.  
RA YOSHIDA K.-I., SEKI S., FUJIMURA M., MIWA Y., YOSHIDA K.-I.;  
RL DNA RES. 2:61-69(1995).  
CC -!- SURCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
DR EMBL; D45242; G90420; -.  
DR SUBTILIS; BG11360; YXBC.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2.  
KW HYPOTHETICAL PROTEIN; TRANSPORT; TRANSMEMBRANE.  
FT TRANSMEM 15 35 POTENTIAL.  
FT TRANSMEM 39 59 POTENTIAL.  
FT TRANSMEM 77 97 POTENTIAL.  
FT TRANSMEM 105 125 POTENTIAL.  
FT TRANSMEM 140 160 POTENTIAL.  
FT TRANSMEM 164 184 POTENTIAL.  
FT TRANSMEM 242 262 POTENTIAL.  
FT TRANSMEM 281 301 POTENTIAL.  
FT TRANSMEM 309 329 POTENTIAL.  
FT TRANSMEM 342 362 POTENTIAL.  
SQ SEQUENCE 388 AA; 42163 MW; 7E9AEDF4 CRC32;  
Query Match 79.4%; Score 50; DB 11; Length 388;  
Best Local Similarity 87.5%; Pred. No. 2.48e+00;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Db 285 gigilnvl 292  
QY 3 GIGILTVI 10  
RESULT 3  
ID YSA1\_YEAST STANDARD; PRT; 231 AA.  
AC Q01976;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE YSA1 PROTEIN.  
GN YSA1 OR YBR111C OR YBR0907.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RX MEDLINE; 92327848;  
RA MANNAHAPT G., STUCKA R., EHNLE S., VETTER I., FELDMANN H.;  
RL YEAST 10:1363-1381(1994).  
RN [2]  
RP SEQUENCE OF 1-47 FROM N.A.  
RC STRAIN-S288C;  
RX MEDLINE; 92327848.  
RA MANNAHAPT G., STUCKA R., EHNLE S., VETTER I., FELDMANN H.;  
RL YEAST 8:397-408(1992).  
CC -!- SIMILARITY: STRONG, TO B.SUBTILIS YKQG.  
CC -!- SIMILARITY: TO PROTEINS WITH A CORE MUTT DOMAIN.  
DR EMBL; Z35980; G536466; -.  
DR EMBL; X78993; G476067; -.  
DR EMBL; X66247; G3549; -.  
DR PIR; S44691; S44691.  
DR LISTA; SC01415; YSA1.  
DR SGD; L0002551; YSA1.  
DR PROSITE; PS00893; MUTT.  
FT DOMAIN 112 145 MUTT-LIKE.  
SQ SEQUENCE 231 AA; 26087 MW; 49A2D6CB CRC32;

Query Match 76.2%; Score 48; DB 11; Length 231;  
Best Local Similarity 75.0%; Pred. No. 6.49e+00;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 79 gigiltil 86  
QY 3 GIGILTVI 10  
RESULT 4  
ID ATPL\_SULAC STANDARD; PRT; 101 AA.  
AC P23040;  
DT 01-NOV-1991 (REL. 20, CREATED)  
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE MEMBRANE-ASSOCIATED ATPASE C CHAIN (EC 3.6.1.34) (SUL-ATPASE  
DE PROTEOLIPID CHAIN).  
GN ATPP.  
OS SULFOLOBUS ACIDOCALDARIUS.  
OC ARCHAEABACTERIA; CRENARCHAEOTA; SULFOLOBALES.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 89214142.  
RA DENDA K., KONISHI J., OSHIMA T., DATE T., YOSHIDA M.;  
RL J. BIOL. CHEM. 264:7119-7121(1989).  
CC -!- FUNCTION: THE C CHAIN IS A PROTEOLIPID, AND ONE OF THE MEMBRANOUS  
CC SUBUNITS OF THE THE NONENZYMATIC COMPONENT OF THE SUL-ATPASE  
CC COMPLEX.  
CC -!- SUBUNIT: SUL-ATPASE IS COMPOSED OF SIX (OR FIVE ?) SUBUNITS:  
CC ALPHA, BETA, DELTA, GAMMA, C (PROTEOLIPID), AND POSSIBLY EPSILON.  
CC -!- SIMILARITY: TO THE PROTEOLIPID SUBUNIT OF FO1-ATPASES.  
DR EMBL; J04740; G152925; -.  
DR PIR; A33351; A33351.  
KW HYDROGEN ION TRANSPORT; LIPID-BINDING; TRANSMEMBRANE.  
FT TRANSMEM 5 25 POTENTIAL.  
FT TRANSMEM 37 57 POTENTIAL.  
FT TRANSMEM 75 95 POTENTIAL.  
SQ SEQUENCE 101 AA; 10362 MW; 1DC8C74D CRC32;  
Query Match 74.6%; Score 47; DB 1; Length 101;  
Best Local Similarity 87.5%; Pred. No. 1.04e+01;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Db 59 aagigvlt 66  
QY 1 AAGIGILT 8  
RESULT 5  
ID MENA\_HAEIN STANDARD; PRT; 308 AA.  
AC P44739;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE 1,4-DIHYDROXY-2-NAPHTHOATE OCTAPRENYLTRANSFERASE (EC 2.5.-.-).  
GN MENA OR HI0509.  
OS HAEMOPHILUS INFLUENZAE.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC PASTEURELLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20;  
RX MEDLINE; 95350630.  
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., KELLEY J.M.,  
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODE A., KELLEY J.M.,  
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,  
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,  
RA GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
RA VENTER J.C.;  
RN SCIENCE 269:496-512(1995).

CC -!- FUNCTION: CONVERSION OF 1,4-DIHYDROXY-2-NAPHTHOATE TO DIMETHYL-  
 CC MENAQUINONE (BY SIMILARITY).  
 CC -!- PATHWAY: MENAQUINONE BIOSYNTHESIS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -!- SIMILARITY: TO B.SUBTILIS YWAB (IPA-6D).  
 DR EMBL; L45150; G1003902; -;  
 DR EMBL; U32732; G925532; -;  
 KW MENAQUINONE BIOSYNTHESIS; TRANSFERASE; TRANSMEMBRANE.  
 FT TRANSMEM 22 42 POTENTIAL.  
 FT TRANSMEM 47 67 POTENTIAL.  
 FT TRANSMEM 101 121 POTENTIAL.  
 FT TRANSMEM 129 149 POTENTIAL.  
 FT TRANSMEM 153 173 POTENTIAL.  
 FT TRANSMEM 186 206 POTENTIAL.  
 FT TRANSMEM 235 255 POTENTIAL.  
 FT TRANSMEM 286 306 POTENTIAL.  
 SQ SEQUENCE 308 AA; 33345 MW; 090B2655 CRC32;  
 Query Match 74.6%; Score 47; DB 6; Length 308;  
 Best Local Similarity 55.8%; Pred. No. 1.04e+01;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 132 agigillaiv 140  
 QY 2 AGIGILTVI 10  
 RESULT 6  
 ID YD88\_SCHPO STANDARD; PRT; 977 AA.  
 AC Q10495;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DE HYPOTHETICAL 111.4 KD PROTEIN C26F1.08C IN CHROMOSOME I.  
 GN SPAC26F1.08C  
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA BROWN D., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;  
 RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 DR EMBL; Z73100; E241983; -;  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 24 44 POTENTIAL.  
 FT TRANSMEM 54 74 POTENTIAL.  
 FT TRANSMEM 83 103 POTENTIAL.  
 FT TRANSMEM 108 128 POTENTIAL.  
 FT TRANSMEM 164 184 POTENTIAL.  
 FT TRANSMEM 376 396 POTENTIAL.  
 FT TRANSMEM 585 605 POTENTIAL.  
 FT TRANSMEM 616 636 POTENTIAL.  
 FT TRANSMEM 656 676 POTENTIAL.  
 FT TRANSMEM 682 702 POTENTIAL.  
 FT TRANSMEM 717 737 POTENTIAL.  
 FT TRANSMEM 857 877 POTENTIAL.  
 FT CARBOHYD 210 210 POTENTIAL.  
 FT CARBOHYD 262 262 POTENTIAL.  
 FT CARBOHYD 265 265 POTENTIAL.  
 FT CARBOHYD 300 300 POTENTIAL.  
 FT CARBOHYD 346 346 POTENTIAL.  
 FT CARBOHYD 364 364 POTENTIAL.  
 FT CARBOHYD 496 496 POTENTIAL.  
 FT CARBOHYD 773 773 POTENTIAL.  
 FT CARBOHYD 926 926 POTENTIAL.  
 FT CARBOHYD 969 969 POTENTIAL.  
 SQ SEQUENCE 977 AA; 111352 MW; C5ACC99F CRC32;  
 Query Match 73.0%; Score 46; DB 11; Length 977;  
 Best Local Similarity 50.0%; Pred. No. 1.65e+01;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 588 aigigllsiv 597  
 QY 1 AAGIGILTVI 10  
 RESULT 7  
 ID NARQ\_ECOLI STANDARD; PRT; 566 AA.  
 AC F27896;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE NITRATE/NITRITE SENSOR PROTEIN NARQ (EC 2.7.3.-).  
 GN NARQ.  
 OS ESCHERICHIA COLI.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE; 92409527.  
 RA RABIN R.S., STEWART V.;  
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:8419-8423(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MC4100;  
 RX MEDLINE; 92374842.  
 RA CHIANG R.C., CAVICCHIOLI R., GUNSALUS R.P.;  
 RL MOL. MICROBIOL. 6:1913-1923(1992).  
 RN [3]  
 RP SEQUENCE OF 547-566 FROM N.A.  
 RC STRAIN-K12;  
 RA NILES M.L., BERTRAND K.P.;  
 RL SUBMITTED (JUL-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: ACTS AS A SENSOR FOR NITRATE/NITRITE AND TRANSDUCES  
 CC SIGNAL OF NITRATE/NITRITE AVAILABILITY TO THE NARL/NARP PROTEINS.  
 CC NARQ PROBABLY ACTIVATES NARL AND NARP BY PHOSPHORYLATION. NARQ  
 CC PROBABLY NEGATIVELY REGULATES THE NARL PROTEIN BY  
 CC DEPHOSPHORYLATION.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
 CC (PROBABLE).  
 CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE  
 CC KINASES.  
 DR EMBL; M94724; G146928; -;  
 DR EMBL; X65714; G581139; -;  
 DR EMBL; U12598; G529012; -;  
 DR EMBL; A46162; A46162.  
 DR EMBL; EGI1450; NARQ.  
 KW SENSORY TRANSDUCTION; TRANSFERASE; KINASE; PHOSPHORYLATION;  
 KW TRANSMEMBRANE; INNER MEMBRANE; NITRATE ASSIMILATION.  
 FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 14 34 POTENTIAL.  
 FT DOMAIN 35 146 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 147 167 POTENTIAL.  
 FT DOMAIN 168 566 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 347 566 TRANSMITTER DOMAIN (POTENTIAL).  
 FT MOD.RES 370 370 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 566 AA; 63696 MW; B4882856 CRC32;  
 Query Match 71.4%; Score 45; DB 6; Length 566;  
 Best Local Similarity 60.0%; Pred. No. 2.59e+01;  
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 156 agigiftlv 165  
 QY 1 AAGIGILTVI 10  
 RESULT 8  
 ID YQHA\_BACSU STANDARD; PRT; 278 AA.  
 AC P54504;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 31.8 KD PROTEIN IN SODA-COMGA INTERGENIC REGION.  
GN YQHA.  
OS BACILLUS SUBTILIS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168 / JH642;  
RA KOBAYASHI Y., MIZUNO M., MASUDA S., TAKEMARU K., HOSONO S.,  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- SIMILARITY: TO B.SUBTILIS YCCR.  
DR EMBL; D84432; G1303874; -.  
DR SUBTILIS; BGL1694; YQHA.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 278 AA; 31810 MW; 858E4F05 CRC32;  
Query Match 69.8%; Score 44; DB 11; Length 278;  
Best Local Similarity 50.0%; Pred. No. 4.05e+01;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
Db 167 tdgigilplv 176  
: ||||| :  
Qy 1 AAGIGILTVI 10  
RESULT 9  
ID GLMU\_ECOLI STANDARD; PRT; 456 AA.  
AC P17114;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE UDP-N-ACETYLGLUCOSAMINE PYROPHOSPHORYLASE (EC 2.7.7.23) (N-  
DE ACETYLGLUCOSAMINE-1-PHOSPHATE URIDYLTRANSFERASE).  
GN GLMU.  
OS ESCHERICHIA COLI.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 85121806.  
RA WALKER J.E., GAY N.J., SARASTE M., EBERLE A.N.;  
RL BIOCHEM. J. 224:799-815(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE; 93315143.  
RA BURLAND V.D., PLUNKETT G. III, DANIELS D.L., BLATTNER F.R.;  
RL GENOMICS 16:551-561(1993).  
RN [3]  
RP IDENTIFICATION.  
RX MEDLINE; 94012475.  
RA MENGIN-LECUREUX D., VAN HEIJENOORT J.;  
RL J. BACTERIOL. 175:6150-6157(1993).  
CC -1- FUNCTION: BIFUNCTIONAL ENZYME RESPONSIBLE FOR THE ACETYLATION OF  
CC GLC-N-1-P TO GIVE GLCNAC-1-P AND THE SYNTHESIS OF UDP-GLCNAC.  
CC -1- CATALYTIC ACTIVITY: UTP + N-ACETYL-ALPHA-D-GLUCOSAMINE  
CC 1-PHOSPHATE = PYROPHOSPHATE + UDP-N-ACETYL-D-GLUCOSAMINE.  
CC -1- PATHWAY: PEPTIDOGLYCAN AND LIPOPOLYSACCHARIDE BIOSYNTHESIS.  
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
CC FRAMESHIFT THAT CREATES TWO ORFS.  
CC -1- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF  
CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).  
DR EMBL; X01631; G43267; -.  
DR EMBL; L10328; G290579; ALT\_FRAME.  
DR EMBL; L10328; G290578; ALT\_FRAME.  
DR ECOGENE; EG11198; GLMU  
DR PROSITE; PS00101; HEXAPEP\_TRANSFERASES.  
KW PEPTIDOGLYCAN SYNTHESIS; CELL WALL; TRANSFERASE;  
KW NUCLEOTIDYLTRANSFERASE; REPEAT; MULTIFUNCTIONAL ENZYME.  
SQ SEQUENCE 456 AA; 49162 MW; 7A80D509 CRC32;  
Query Match 69.8%; Score 44; DB 4; Length 456;  
Best Local Similarity 75.0%; Pred. No. 4.05e+01;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 124 ggigilty 131  
: ||||| :  
Qy 2 AGIGILTV 9  
RESULT 10  
ID FIMB\_DICDI STANDARD; PRT; 610 AA.  
AC P54680;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE FIMBRIN.  
GN FIMA.  
OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).  
OC EUKARYOTA; PROTOZOA; SARCOMASTIGOPHORA; SARCODINA; RHIZOPODA;  
OC EUMYCETOZOA; DICTYOSTELIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AX3;  
RA PRASSLER J.;  
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- FUNCTION: BINDS TO ACTIN.  
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY  
CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,  
CC ABP-120, ABP-180, OR BETA-FODRIN).  
CC EMBL; L36202; G992560; -.  
DR DICTYDB; DD0777; FIMA.  
KW CALCIUM-BINDING; ACTIN-BINDING; DUPLICATION.  
FT CA\_BIND 20 31 POTENTIAL.  
FT CA\_BIND 56 67 POTENTIAL.  
FT DOMAIN 102 363 ACTIN-BINDING 1.  
FT DOMAIN 364 604 ACTIN-BINDING 2.  
SQ SEQUENCE 610 AA; 66975 MW; 6B522C77 CRC32;  
Query Match 69.8%; Score 44; DB 4; Length 610;  
Best Local Similarity 60.0%; Pred. No. 4.05e+01;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Db 532 stgipildvi 541  
: ||||| :  
Qy 1 AAGIGILTVI 10  
RESULT 11  
ID YD3M\_HERAU STANDARD; PRT; 611 AA.  
AC P25280;  
DT 01-MAY-1992 (REL. 22, CREATED)  
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 68.4 KD PROTEIN IN HGDIIM 3'REGION (ORF68).  
OS HERPETOSEPHON AURANTIACUS (HERPETOSEPHON GIGANTEUS).  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA;  
OC NONPHOTOSYNTHETIC, NONFRUITING GLIDING; BEGGIATOACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HPA2;  
RX MEDLINE; 92039068.  
RA DUBSTERHOEFT A., KROEGER M.;  
RL GENE 106:87-92(1991).  
DR EMBL; X55141; G48774; -.  
DR PIR; JT0592; JT0592.  
DR PIR; S21953; S21953.  
DR PIR; S21950; S21950.  
KW HYPOTHETICAL PROTEIN; RESTRICTION SYSTEM; REPEAT.  
FT DOMAIN 382 403 2.5 X 11 AA TANDEM REPEATS.  
FT REPEAT 382 392 1.  
FT REPEAT 393 403 2.  
FT REPEAT 404 409 3 (INCOMPLETE).  
SQ SEQUENCE 611 AA; 68354 MW; 473CD6A4 CRC32;

```

Query Match          69.8%; Score 44; DB 11; Length 611;
Best Local Similarity 62.5%; Pred. No. 4.05e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 97 gigilaif 104
QY 3 GIGILTVI 10

RESULT 12
ID XYND_BACPO STANDARD; PRT; 635 AA.
AC P45796;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE ENDO-1.4-BETA-XYLANASE D PRECURSOR (EC 3.2.1.8) (XYLANASE D)
DE (1.4-BETA-D-XLAN XYLANOHYDROLASE D).
GN XYND.
OS BACILLUS POLYMYXA.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 842;
RX MEDLINE; 92041687.
RA GOSALBES M.J., PEREZ-GONZALEZ J.A., GONZALEZ R., NAVARRO A.;
RL J. BACTERIOL. 173:7705-7710(1991).
CC -!- FUNCTION: SHOWS XYLANASE ACTIVITY AS WELL AS ALPHA-L-
CC ARABINOFURANOSIDASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC LINKAGES IN XYLANS.
CC -!- PATHWAY: XLAN DEGRADATION.
CC -!- SIMILARITY: BELONGS TO FAMILY 43 OF GLYCOSYL HYDROLASES.
DR EMBL; X57094; G48816; -.
KW XLAN DEGRADATION; HYDROLASE; GLYCOSIDASE; SIGNAL.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 635 XYLANASE D.
SQ SEQUENCE 635 AA; 67914 MW; 078AAB82 CRC32;

Query Match          69.8%; Score 44; DB 10; Length 635;
Best Local Similarity 75.0%; Pred. No. 4.05e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 149 gagigvlt 156
QY 1 AAGIGILT 8

RESULT 13
ID NCAL_CHICK STANDARD; PRT; 1091 AA.
AC P13590;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE NEURAL CELL ADHESION MOLECULE, LARGE ISOFORM PRECURSOR (N-CAM 180)
DE (CONTAINS: N-CAM 140).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN RP SEQUENCE OF 1-175 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE; 87206190.
RA CUNNINGHAM B.A., HEMPERLY J.J., MURRAY B.A., PREDIGER E.A.,
RA BRACKENBURY R., EDELMAN G.M.;
RL SCIENCE 236:799-806(1987).
[2]
RN RP SEQUENCE OF 128-1091 FROM N.A., AND PARTIAL SEQUENCE.
RX HEMPERLY J.J., MURRAY B.A., EDELMAN G.M., CUNNINGHAM B.A.;
RA HEMPERLY J.J., MURRAY B.A., EDELMAN G.M., CUNNINGHAM B.A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 83:3037-3041(1986).
CC -!- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN
CC NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF
CC NEURITES, ETC.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

```

```

CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC FIVE C2-LIKE DOMAINS FOLLOWED BY 2 FIBRONECTIN TYPE III-LIKE
CC DOMAINS.
CC -!- ALTERNATIVE PRODUCTS: THE DIFFERENT TISSUE-SPECIFIC FORMS OF
CC N-CAM ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
DR EMBL; M15861; G212436; ALT_SEQ.
DR EMBL; M15860; G212436; JOINED.
DR EMBL; M15922; G212436; JOINED.
DR EMBL; M15923; G212436; JOINED.
DR EMBL; M15924; G212436; JOINED.
DR EMBL; M21178; G212436; JOINED.
DR EMBL; M21179; G212436; JOINED.
DR EMBL; M21180; G212436; JOINED.
DR EMBL; M15929; G212436; JOINED.
DR EMBL; M15930; G212436; JOINED.
DR EMBL; M15931; G212436; JOINED.
DR EMBL; M15932; G212436; JOINED.
DR EMBL; M15933; G212436; JOINED.
DR EMBL; M15934; G212436; JOINED.
DR EMBL; L29437; G212436; JOINED.
DR EMBL; M15935; G212436; JOINED.
DR EMBL; M15937; G212436; JOINED.
DR EMBL; M15938; G212436; JOINED.
DR EMBL; M15939; G212436; JOINED.
DR PIR; A43613; IUCHNL.
KW CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW IMMUNOGLOBULIN FOLD; ALTERNATIVE SPLICING; SIGNAL.
FT SIGNAL 1 19
FT CHAIN 20 1091 NEURAL CELL ADHESION MOLECULE, LARGE
FT ISOFORM.
FT DOMAIN 20 711 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 712 729 POTENTIAL.
FT DOMAIN 730 1091 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 34 103 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 132 196 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 228 294 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 322 392 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 419 486 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 518 595 FIBRONECTIN TYPE-III.
FT DOMAIN 624 692 FIBRONECTIN TYPE-III.
FT DOMAIN 152 165 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 161 165 BY SIMILARITY.
FT DISULFID 41 96 BY SIMILARITY.
FT DISULFID 139 189 BY SIMILARITY.
FT DISULFID 235 287 BY SIMILARITY.
FT DISULFID 329 385 BY SIMILARITY.
FT DISULFID 426 479 BY SIMILARITY.
FT CARBOHYD 222 222 POTENTIAL.
FT CARBOHYD 315 315 POTENTIAL.
FT CARBOHYD 347 347 POTENTIAL.
FT CARBOHYD 423 423 POTENTIAL.
FT CARBOHYD 449 449 POTENTIAL.
FT CARBOHYD 478 478 POTENTIAL.
FT CARBOHYD 800 800 POTENTIAL.
FT CARBOHYD 914 914 POTENTIAL.
FT CARBOHYD 945 945 POTENTIAL.
FT CARBOHYD 954 954 POTENTIAL.
FT VARSPIC 810 1070 MISSING (IN SHORT N-CAM 140 FORM).
SQ SEQUENCE 1091 AA; 117397 MW; A97DFEEL CRC32;

Query Match          69.8%; Score 44; DB 6; Length 1091;
Best Local Similarity 70.0%; Pred. No. 4.05e+01;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 712 aalvgilivi 721
QY 1 AAGIGILTVI 10

RESULT 14
ID YHAH_ECOLI STANDARD; PRT; 130 AA.
AC P42621;
DT 01-NOV-1995 (REL. 32, CREATED)

```

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 15.4 KD PROTEIN IN EXOR-TDCC INTERGENIC REGION (OI30).  
GN YHAH.  
OS ESCHERICHIA COLI.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RA PLUNKETT G. III;  
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

DR EMBL; U18997; G606044; -.  
DR ECOGENE; EGI2747; YHAH.  
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
FT TRANSMEM 24 44 POTENTIAL.  
FT TRANSMEM 50 70 POTENTIAL.  
FT TRANSMEM 81 101 POTENTIAL.  
SQ SEQUENCE 130 AA; 15370 MW; DC181FCD CRC32;

Query Match 68.3%; Score 43; DB 11; Length 130;  
Best Local Similarity 70.0%; Pred. No. 6.27e+01;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 49 aggegiltti 58  
| | | | |  
QY 1 AAGIGILTVI 10

RESULT 15  
ID ATPE\_ARATH STANDARD; PRT; 132 AA.  
AC P09468;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)  
DE 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
DE ATP SYNTHASE EPSILON CHAIN (EC 3.6.1.34).  
GN ATPE.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OG CHLOROPLAST  
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
OC CAPPARALES; CRUCIFERAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. LANDSBERG ERECTA;  
RX MEDLINE; 89057486.  
RA CHEN H.-C.; WINTZ H.; WEIL J.-H.; PILLAY D.T.N.;  
RL NUCLEIC ACIDS RES. 16:10372-10372(1988).  
CC -I- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
GRADIENT ACROSS THE MEMBRANE.  
CC -I- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS. CF(1) - THE CATALYTIC  
CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
HAS THREE MAIN SUBUNITS: A, B AND C.  
CC -I- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.  
CC -I- SIMILARITY: BELONGS TO THE ATPASE EPSILON CHAIN FAMILY.  
DR EMBL; X12889; G11334; -.  
PIR; S01903; S01903.  
KW ATP SYNTHESIS; CHLOROPLAST; THYLAKOID MEMBRANE; CF(1);  
KW HYDROLASE; HYDROGEN ION TRANSPORT.  
SQ SEQUENCE 132 AA; 14472 MW; D826F274 CRC32;

Query Match 68.3%; Score 43; DB 1; Length 132;  
Best Local Similarity 66.7%; Pred. No. 6.27e+01;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 43 avdigilti 51  
| | | | |  
QY 1 AAGIGILTV 9

Search completed: Tue Jun 10 11:19:55 1997  
Job time : 9 secs.

\*\*\*\*\*  
WQSEPH  
\*\*\*\*\* (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:22:06-1997; MasPar time 1.79 Seconds  
60.838 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-231-565A-31  
Description: (1-10) from US08231565A.pep  
Perfect Score: 68  
Sequence: 1 LLDGTATLRL 10

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 08  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 16.764; Variance 52.447; scale 0.320  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	68	100.0	10	14	R78639 Immunogenic peptide o	6.70e-01
2	68	100.0	10	15	R84199 gp100 melanoma antige	6.70e-01
3	68	100.0	10	15	R82167 Melanoma-specific mut	6.70e-01
4	68	100.0	661	14	R78646 Melanoma associated a	6.70e-01
5	68	100.0	661	15	R84855 MART-1 melanoma antig	6.70e-01
6	68	100.0	661	15	R84854 MART-1 melanoma antig	6.70e-01
7	51	75.0	9	15	R82124 Melanoma-specific mut	4.72e-01
8	50	73.5	766	11	R59925 GAP protein Gapi.	5.99e-01
9	48	70.6	318	3	R20110 Streptomyces clavulig	9.63e-01
10	48	70.6	338	3	R21420 Streptomyces clavulig	9.63e-01
11	48	70.6	3398	8	R44430 eRYA region polypepti	9.63e-01
12	45	66.2	589	3	P60303 Sequence encoded by t	1.94e-02
13	45	66.2	1684	3	R14948 Bacterial amylase A-1	1.94e-02
14	44	64.7	16	13	R68954 Wild type interleukin	2.44e-02
15	44	64.7	158	15	R77392 Human mutant IL-6a C2	2.44e-02
16	44	64.7	181	19	W05586 -4aa SSCC mutein of h	2.44e-02
17	44	64.7	181	6	R31996 -4aa IL-6 SSSS mutein	2.44e-02
18	44	64.7	182	14	R75346 Hybrid human cytokine	2.44e-02
19	44	64.7	184	1	P81158 Polypeptide with B-ce	2.44e-02
20	44	64.7	184	12	R72219 Human interleukin-6 -	2.44e-02

21	44	64.7	184	3	R03914 Polypeptide with huma	2.44e+02
22	44	64.7	184	1	R05895 Human B-cell differen	2.44e+02
23	44	64.7	185	12	R72215 Human interleukin-6 (	2.44e+02
24	44	64.7	185	13	R68624 Ala-BCDF.	2.44e+02
25	44	64.7	185	2	R10983 Recombinant human B-c	2.44e+02
26	44	64.7	185	12	R72220 Human interleukin-6 (	2.44e+02
27	44	64.7	185	1	P90059 Human Ala B cell diff	2.44e+02
28	44	64.7	185	9	R45718 Full length interleuk	2.44e+02
29	44	64.7	185	2	R05311 Segment of B-cell dif	2.44e+02
30	44	64.7	186	19	W05585 SSCC mutein of human	2.44e+02
31	44	64.7	186	19	W05587 SSCC mutein of human	2.44e+02
32	44	64.7	186	14	R75343 Hybrid human cytokine	2.44e+02
33	44	64.7	200	19	R05588 16 aa sig -1 aa SSCC	2.44e+02
34	44	64.7	208	2	P94755 Sequence of variant o	2.44e+02
35	44	64.7	211	5	R25279 Mutant human BCDF.	2.44e+02
36	44	64.7	212	7	R33384 Cytokine hIL-6.	2.44e+02
37	44	64.7	212	13	R72317 Interferon-beta2A.	2.44e+02
38	44	64.7	212	1	P90121 Human lymphocyte rece	2.44e+02
39	44	64.7	212	1	P90469 Interleukin-6	2.44e+02
40	44	64.7	212	3	R81176 Sequence of the BSF-2	2.44e+02
41	44	64.7	212	9	R49249 Sequence of human B-c	2.44e+02
42	44	64.7	212	1	P90047 pBSF2-L8 sequence	2.44e+02
43	44	64.7	212	1	R05415 Human B-cell differen	2.44e+02
44	44	64.7	258	1	R05777 Cysteine-free interie	2.44e+02
45	44	64.7	310	2	R03255 Fusion protein of B-	2.44e+02

ALIGNMENTS

RESULT 1  
ID R78639 standard; Protein; 10 AA.  
AC R78639;  
DT 22-JAN-1996 (first entry)  
DE Immunogenic peptide of melanoma associated antigen gp100.  
KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
KW Identification; tumour; gp100.  
OS Homo sapiens.  
PN EP-668390-A1  
PD 23-AUG-1995  
PF 14-FEB-1995; 200348.  
PR 16-FEB-1994; EP-200337.  
PR 21-DEC-1994; EP-203709.  
PA (ALKU AKU NOBEL NV.  
PI Adema Gv. Fiqdor CG;  
DR WPI; 95-284790/38.  
DR N-PSDB; Q96055.  
PT Melanoma associated antigen gp100 - used in vaccines and for the  
PT Detection of tumours  
PS Claim 5; Page 31; 40pp; English.  
CC Immunogenic peptides derived from the melanoma associated antigen  
CC (See R78639-45) may be used in the production of vaccines.  
CC Nucleotide sequences encoding the immunogenic peptides may be used  
CC as primers and probes in the detection of melanoma cells. Tumour  
CC infiltrating lymphocytes capable of binding to the melanoma  
CC associated antigen can be cultured ex vivo and returned to melanoma  
CC particles, and when radiolabelled, they may be used to identify  
CC tumour deposits.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 68; DB 14; Length 10;  
Best Local Similarity 100.0%; Pred. No. 6.70e-01;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 lldgtatlrl 10  
|||||  
QY 1 LLDGTATLRL 10

RESULT 2  
ID R84199 standard; Peptide; 10 AA.  
AC R84199;  
DT 20-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G10-4).

KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN W09523093-A2.

PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 55; Page 131; 184pp; English.  
 CC The immunogenic peptide is derived from CDNA25 (R84854), a  
 CC melanoma antigen derivative of gp100 (see R84855). The  
 CC peptide and its derivatives (see R84200-R84211) are used in  
 CC medicaments (vaccines) for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against melanoma-specific  
 CC antigens and its immunogenic peptides may be used in the  
 CC detection and isolation of the antigen from a sample, the  
 CC detection of which is indicative of a disease state  
 CC (melanoma or metastatic melanoma).  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 68; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.70e-01;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 lldgtatlrl 10  
 |||||  
 Qy 1 LLDGTATLRL 10

RESULT 3  
 ID R82167 standard; peptide; 10 AA.  
 AC R82167;  
 DT 25-MAR-1996 (first entry)  
 DE Melanoma-specific mutant immunogen epitope 10mer peptide.  
 DE Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
 KW cytotoxic T cell; lymphocyte; HLA-A2.  
 OS Homo sapiens.  
 PN W09523561-A2.  
 PD 24-AUG-1995.  
 PF 16-FEB-1995; U01991.  
 PR 16-FEB-1994; US-197399.  
 PR 29-APR-1994; US-234784.  
 PA (UNIV-) UNIV VIRGINIA PATENT FOUND.  
 PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
 DR WPI; 95-302688/39.  
 PT Melanoma-specific immunogen comprises epitope(s) homologous with  
 PT pMEL-17 - are highly potent stimulators of HLA-A2+CTL's useful in  
 PT adoptive immuno-therapy  
 PS Example 8; Page 52; 148pp; English.  
 CC A melanoma-specific immunogen homologous with pMEL-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
 CC R82108 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pMEL-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 68; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.70e-01;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 lldgtatlrl 10  
 |||||

Qy 1 LLDGTATLRL 10

RESULT 4  
 ID R78646 standard; Protein; 661 AA.  
 AC R78646;  
 DT 22-JAN-1996 (first entry)  
 DE Melanoma associated antigen gp100.  
 DE Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 PN EP-669350-A1.  
 PD 23-AUG-1995.  
 PF 14-FEB-1995; 200348  
 PR 16-FEB-1994; EP-200337.  
 PR 21-DEC-1994; EP-203709.  
 PA (ALKU ) AKZO NOBEL NV.  
 PI Adema GJ, Figdor CG;  
 DR WPI; 95-284790/38.  
 DR N-PSDB; Q96055.  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PT detection of tumours  
 PS Claim 1; Page 22-24; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC may be used in the production of vaccines. Nucleotide sequences  
 CC encoding the immunogenic peptides may be used as primers and probes  
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
 CC capable of binding to the melanoma associated antigen can be  
 CC cultured ex vivo and returned to melanoma particles, and when  
 CC radiolabelled, they may be used to identify tumour deposits.  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 68; DB 14; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 6.70e-01;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 457 lldgtatlrl 466  
 |||||  
 Qy 1 LLDGTATLRL 10

RESULT 5  
 ID R84855 standard; Protein; 661 AA.  
 AC R84855;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen gp100.  
 KW gp100; MART-1; melanoma antigen recognised by T-cell;  
 KW CDNA25 antigen derivative; melanocyte; melanoma; immunogen;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian sp.  
 FH Key Location/Qualifiers  
 FT Peptide 154..163  
 FT /label= G9-154\_immunogenic\_peptide  
 FT Peptide 208..217  
 FT /label= G9-209\_immunogenic\_peptide  
 FT /note= "see R84210"  
 FT Peptide 280..288  
 FT /label= G9-280\_immunogenic\_peptide  
 FT /note= "see R84208"  
 FT Peptide 457..266  
 FT /label= Immunogenic\_peptide  
 FT Peptide 476..485  
 FT /label= Immunogenic\_peptide  
 PN W09529198-A2.  
 PD 02-NOV-1995.  
 PR 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 7A; 184pp; English.  
 CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which  
 CC is recognized by T-lymphocytes, and is a derivative of the  
 CC melanoma-specific antigen CDNA25 (see R84854). gp100 is a source  
 CC of immunogenic peptides which are optionally modified to enhance  
 CC their binding to a MHC molecule, and used in medicaments,  
 CC especially vaccines, for the treatment or prevention (by  
 CC immunisation) of melanoma. Antibodies against CDNA2 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 68; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 6.70e-01;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 457 lldgtatlrl 466  
 |||||  
 QY 1 LLDGTATLRL 10

## RESULT 6

ID R84854 standard; Protein; 661 AA.

AC R84854;  
 DT 08-MAY-1996 (revised)  
 DE 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen CDNA25.  
 KW CDNA25; melanoma antigen recognised by T-cell;  
 KW gp100 antigen derivative; melanoma; metastatic melanoma;  
 KW tumour-associated antigen; immunogen; diagnosis; prognosis;  
 KW prophylaxis; therapy; vaccine.  
 OS Mammalian sp.

FH Key Location/Qualifiers

FT Peptide 457..466  
 FT //label antigenic-peptide  
 FT //note= see R84199"  
 PN W0952919-A2.

PD 02-NOV-1995.

PF 31-APR-1995; U05063.

PR 28-APR-1994; US-231565.

PR 05-APR-1995; US-417174.

PA (USSH ) US SEC DEPT HEALTH.

PI Kawakami Y, Rosenberg SA;

DR WPI; 95-382963/49.

DR N-PSDB; T02716.

PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.

PS Claim 81; Fig 5A; 184pp; English.

CC CDNA2 is a melanoma antigen (MART-1) which is recognized by  
 CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-  
 CC specific antigen gp100 (see R84855). Antigen CDNA25 is a source  
 CC of immunogenic peptides (see R84199) which are optionally modified  
 CC (see R84200-R84211) to enhance their binding to a MHC molecule and  
 CC used in medicaments, especially vaccines, for the treatment or  
 CC prevention (by immunisation) of melanoma. Antibodies against CDNA2  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of the antigen from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 68; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 6.70e-01;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 457 lldgtatlrl 466  
 |||||  
 QY 1 LLDGTATLRL 10

## RESULT 7

ID R82124 standard; peptide; 9 AA.

AC R82124;

DT 25-MAR-1996 (first entry)

DE Melanoma-specific mutant immunogen epitope 9mer peptide.

KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;

KW Cytotoxic T cell; lymphocyte; HLA-A2.

OS Homo sapiens.

PN W09522561-A2.

PD 24-JUG-1995.

PF 16-FEB-1995; U01991.

PR 16-FEB-1994; US-197399.

PR 29-APR-1994; US-234784.

PA (UYVI-) UNIV VIRGINIA PATENT FOUND.

PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;

DR WPI; 95-302688/39.

PT Melanoma-specific immunogen comprises epitope(s) homologous with

PT pMel-17 - are highly potent stimulators of HLA-A2+CTL's useful in

PT adoptive immuno-therapy

PS Example 8; Page 51; 148pp; English.

CC A melanoma-specific immunogen homologous with pMel-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
 CC R82108 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pMel-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.  
 SQ Sequence 9 AA;

Query Match 75.0%; Score 51; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.72e+01;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 lldgtatl 9  
 |||||  
 QY 1 LLDGTATL 8

## RESULT 8

ID R59925 standard; protein; 766 AA.

AC R59925;

DT 22-FEB-1995 (first entry)

DE GAP protein Gap1.

KW Ras; GTPase activating protein; GAP; GAP related domain; GRD;

KW Ras; v-Ras; Heat shock; neurofibromatosis type 1; NF1.

OS Schizosaccharomyces pombe.

PN W09416369-A.

PD 24-JUL-1994.

PF 19-JAN-1994; U00198.

PR 15-JAN-1993; US-004824.

PA (SCRE) SCHERING CORP.

PI Kaziro Y, Nakafuku M;

DR WPI; 94-249215/30.

PT Blocking Ras-induced effects on a cell - by introducing a GTPase

PT activating protein to the cell, used esp. in treatment of cancers

PS Disclosure; Page 75-78; 87pp; English.

CC Human neurofibromatosis type 1 (NF1)-GAP related domain (GRD)

CC mutant clones NF201 (given in R59221) and NF204 (R59922) show

CC strong suppression activity for RAS2Val19, and inhibit

CC v-Ras-induced transformation in mammalian cells. The mutation

CC sites of these proteins were located in one of the most conserved

CC regions of GRD. These sites were compared with those of other

CC GRD family proteins, yeast Ira2 (R59926) and Ira1 (R59923),

CC human GAP (R59924) and Schizosaccharomyces pombe Gap1 (R59925).

CC Sequence 766 AA;

Query Match 73.5%; Score 50; DB 11; Length 766;  
 Best Local Similarity 70.0%; Pred. No. 5.99e+01;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 565 lldelstlrl 574



```

CC frame and a completely different protein to the sequence given
CC here. The sequence shown here is identical to that given in the
CC specification, AAs 141-160 appear to have been mistakenly duplicated
CC as there are no corresponding codons for such a duplication in the
CC given DNA sequence. See also R20110.
SQ Sequence 338 AA;

Query Match 70.6%; Score 48; DB 3; Length 338;
Best Local Similarity 60.0%; Pred. No. 9.63e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 176 lldadpvlrl 185
   |||: |||
QY 1 LLDGTATLRL 10

RESULT 11
ID R44430 standard; Protein; 3398 AA.
AC R44430;
DT 22-DEC-1993 (first entry)
DE eryA region polypeptide module #1.
KW Saccharopolyspora erythraea; eryA; biosynthesis; polyketide; module;
KW erythromycin; condensation; elongation; acyl chain growth;
KW gene replacement.
OS Saccharopolyspora erythraea.
PN W09313663-A.
PD 22-JUL-1993.
PF 17-JAN-1992; U00427.
PR 17-JAN-1992; WO-U00427.
PA (ABBO ) ABBOTT LAB.
PI Donadio S, Katz L, McAlpine JB;
DR WPI; 93-242804/30.
DR N-PSDB; Q46806.
PT Biosynthesis of specific polyketide analogues esp. erythromycin
PT cpds. - by introducing altered biosynthetic gene-contg. DNA into
PT microorganisms
PS Disclosure; Fig 2; 133pp; English.
CC The sequences given in R44430-32 are encoded by the eryA fragment of
CC the Saccharopolyspora erythraea genome. These polypeptides are
CC involved in the biosynthesis of the polyketide segment of erythromycin.
CC eryA is organised in modules and each module takes care of one
CC condensation step. The precise succession of elongation steps is
CC dictated by the genetic order of the modules. The DNA encoding
CC these polypeptides may be specifically altered such that novel
CC polyketide molecules of desired structure are produced. Three types
CC of alteration may be produced; those inactivating a single function in
CC a module which does not arrest acyl chain growth; those inactivating a
CC single function in a module which does affect chain growth; and those
CC affecting an entire module. The mutations may be introduced by gene
CC replacement.
SQ Sequence 3398 AA;

Query Match 70.6%; Score 48; DB 8; Length 3398;
Best Local Similarity 50.0%; Pred. No. 9.63e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 3283 llnaatglrl 3292
   ||::: |||
QY 1 LLDGTATLRL 10

RESULT 12
ID P60303 standard; Protein; 589 AA.
AC P60303;
DT 13-SEP-1991 (first entry)
DE Sequence encoded by the breakpoint cluster region (bcr) of
DE chromosome 22 on pvi-3.
KW Chronic myelocytic leukaemia; acute lymphocytic leukaemia;
KW diagnosis; chromosomal translocation; Philadelphia translocation.
OS Homo sapiens.
PN EP-181635-A.
PD 21-MAY-1986.
PF 13-NOV-1985; 114436.

```

---

```

QY 1 LLDGTATLRL 10
   |||: |||

RESULT 9
ID R20110 standard; Protein; 318 AA.
AC R20110;
DT 07-APR-1992 (first entry)
DE Streptomyces clavuligerus hydroxylase.
KW Deacetylcephalosporin C; DAOC; deacetylcephalosporin C; DAC.
OS Streptomyces clavuligerus.
PN EP-465189-A.
PD 08-JAN-1992.
PF 01-JUL-1991; 305939.
PR 06-JUL-1990; US-549502.
PA (ELIL ) ELI LILLY & CO.
PI Kovacevic S, Miller JR;
DR WPI; 92-010373/02.
DR N-PSDB; Q20267.
PT DNA encoding Streptomyces clavuligerus hydroxylase - used for
PT increasing the yield of cephalosporin cpds. and producing new
PT cephalosporin(s)
PS Disclosure; Page 4; 15pp; English.
CC The amino acid (AA) sequence is that of a hydroxylase from
CC Streptomyces clavuligerus which catalyses the reaction in which
CC deacetylcephalosporin C (DAOC) is hydroxylated at the 3-methyl gp.
CC to form deacetylcephalosporin C (DAC). The AA sequence differs
CC from that which would result if the nucleotide sequence given in
CC the specification was translated. It appears that a nucleotide (T)
CC has been omitted from position 307 resulting in a shift of reading
CC frame and a completely different protein to the sequence given in
CC the specification (this also appears to have AAs 141-160 duplicated).
CC See also R21420.
SQ Sequence 318 AA;

Query Match 70.6%; Score 48; DB 3; Length 318;
Best Local Similarity 60.0%; Pred. No. 9.63e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 156 lldadpvlrl 165
   |||: |||
QY 1 LLDGTATLRL 10

RESULT 10
ID R21420 standard; Protein; 338 AA.
AC R21420;
DT 07-APR-1992 (first entry)
DE Streptomyces clavuligerus hydroxylase.
KW Deacetylcephalosporin C; DAOC; deacetylcephalosporin C; DAC.
OS Streptomyces clavuligerus.
FH Key Location/Qualifiers
FT Region 161..180
FT /note= "duplication of AAs 141-160, error"
PN EP-465189-A.
PD 08-JAN-1992.
PF 01-JUL-1991; 305939.
PR 06-JUL-1990; US-549502.
PA (ELIL ) ELI LILLY & CO.
PI Kovacevic S, Miller JR;
DR WPI; 92-010373/02.
DR N-PSDB; Q20267.
PT DNA encoding Streptomyces clavuligerus hydroxylase - used for
PT increasing the yield of cephalosporin cpds. and producing new
PT cephalosporin(s)
PS Disclosure; Page 4; 15pp; English.
CC The amino acid (AA) sequence is that of a hydroxylase from
CC Streptomyces clavuligerus which catalyses the reaction in which
CC deacetylcephalosporin C (DAOC) is hydroxylated at the 3-methyl gp.
CC to form deacetylcephalosporin C (DAC). The AA sequence differs
CC from that which would result if the nucleotide sequence given in
CC the specification was translated. It appears that a nucleotide (T)
CC has been omitted from position 307 resulting in a shift of reading

```

```

PR 14-NOV-1984; US-671296.
PR 26-JUN-1985; US-749178.
PA (ONCO-) ONCOGENE SCI INC.
PI Groffen J, Heisterkamp N, Stephenson JR;
DR WPI; 86-132554/21.
DR N-PSDB; N60228.
PT Detecting chromosomal translocations with single stranded nucleic
PT acid - for detecting human cancer e.g. chronic myelocytic
PT leukaemia
PS Disclosure; Fig 3; 67pp; English.
CC 4 of the 13 exons which pVI-3 (22) encompasses are within the
CC breakpoint cluster region (bcr) (see N60228). To exactly determine
CC the location of the breakpoint in K-562 (22q-), the DNA sequence of
CC the breakpoint region of K-562 was compared with the normal genomic
CC chromosome 22 DNA sequences (see N60229,N60230). In the DNA of
CC patient 0319129, the chromosomal break has occurred in a rather
CC "precise" manner, leading to the generation of a 22q- and 9q+
CC sequence exactly reflecting the sequence of the normal chromosome 22
CC and 9 DNA sequences (N60231,N60232,N61202,N61203). However, in the
CC 9q+ DNA of patient 0212015, sequences are found between the
CC breakpoints on chromosome 9 and 22 which are not present in the
CC region sequenced of the normal chromosomes 9 and 22 (N60233,N60234,
CC N60235).
SQ Sequence 589 AA;

Query Match 66.2%; Score 45; DB 3; Length 589;
Best Local Similarity 55.6%; Pred. No. 1.94e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 284 legsqtlri 292
   i:: |||
Qy 2 LDGTATLRL 10

RESULT 13
ID R14948 standard; Protein; 1684 AA.
AC R14948;
DT 25-FEB-1992 (first entry)
DE Bacterial amylase A-180.
KW Maltopentose; G5; starch hydrolysis.
PN EP-453385-A.
PD 04-DEC-1991.
PF 28-MAY-1991; 108669.
PR 31-MAY-1990; DE-017595.
PA (CONE ) CONSORT ELEKTROCHEM IND.
PI Schmid G, Candussio A, Bock A.
DR WPI; 91-355676/49.
DR N-PSDB; Q13939.
PT New bacterial amylase, A-180 for malto:pentose prodn. - by
PT hydrolysis of starch, providing high yield and modifiable for
PT secretion from host cells
PS Disclosure; Page 7; 21pp; German.
CC The amino acid sequence is that of bacterial amylase A-180 which is
CC used to prepare maltopentose (G5) in high yields by hydrolysis of
CC starch. It can be modified to ensure its excretion, obviating the need
CC to concentrate and purify the enzyme, i.e. the culture supernatant can
CC be used directly for G5 prodn. G5 yields of over 90% are possible,
CC eliminating the need for further G5 purification. G5 is used in medical
CC assays for diagnosis, as carbohydrate sources in liquid feeding compsns.
CC and for converting (by esterification) fatty acids to water-soluble
CC form, suitable for use in stable infusion solns.
SQ Sequence 1684 AA;

Query Match 66.2%; Score 45; DB 3; Length 1684;
Best Local Similarity 77.8%; Pred. No. 1.94e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 564 fdetatrlr 672
   i: |||||
Qy 2 LDGTATLRL 10

RESULT 14

```

```

ID R68954 standard; Peptide; 16 AA.
AC R68954;
DT 05-SEP-1995 (first entry)
DE Wild type interleukin-6 helix A.
KW Human; interleukin-6; IL-6; helix A; biological activity; superagonist;
KW binding activity; receptor; gp80; antagonist; superantagonist; hormone;
KW gp130; receptor complex; oncostatin; leukaemia inhibitory factor; LIF;
KW ciliary neurotrophic factor; CNTF; agonist.
OS Homo sapiens.
PN WO9500852-A.
PD 05-JAN-1995.
PF 23-JUN-1994; IT0095.
PR 23-JUN-1993; IT-RN0409.
PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
PI Ciliberto G, Lahm A, Savino R;
DR WPI; 95-052231/07.
PT Selecting super-agonists, antagonists and super-antagonists or
PT hormones - partic. interleukin-6, oncostatin M, leukaemia
PT inhibitory factor, ciliary neurotrophic factor or interleukin-11
PS Example 2; Page 12; 35pp; English.
CC The amino acid sequence of the helix A of the human interleukin-6 (IL-6).
CC The gene encoding the IL-6 was mutated by PCR to change certain of the
CC binding activity to receptor gp80 for the wild type and mutants sequences
CC could then be compared. The interaction of the IL-6 and its mutants
CC forms part of a method to detect the superagonists, antagonists or
CC superantagonists of a hormone. This is performed by comparing the amino
CC acid sequence of the hormone with its (ant)agonist; comparing the a.a.
CC sequences of the (ant)agonist receptor with the hormone-specific receptor
CC and gp130, and identifying residues forming a part of the site of
CC interaction with the hormone receptor and gp130 based on a 3-dimensional
CC model of the receptor complex. The method can be used to select
CC super(ant)agonists of hormones such as interleukin-6 (IL-6),
CC oncostatin M (OSM) leukaemia inhibitory factor (LIF) or ciliary
CC neurotrophic factor (CNTF).
SQ Sequence 16 AA;

Query Match 64.7%; Score 44; DB 13; Length 16;
Best Local Similarity 55.6%; Pred. No. 2.44e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 6 lldgisalr 14
   :||| :||
Qy 1 LLDGTATLRL 9

RESULT 15
ID R77392 standard; Protein; 158 AA.
AC R77392;
DT 17-APR-1996 (first entry)
DE Human mutant IL-6a/C2 (amino acids 5-19 and 73-83 deleted).
KW Human interleukin-6; IL-6a/C2 mutant; increased stability;
KW recombinant; production; deletion mutant;
KW amino acids 5-19 and 73-83.
OS Homo sapiens.
FH key Location/Qualifiers
FT mat_peptide 1..474
FT /*tag= a
   J07224097-A.
PD 22-AUG-1995.
PF 08-FEB-1994; 014461.
PR 08-FEB-1994; JP-014461.
PA (ASAG ) ASahi GLASS CO LTD.
DR WPI; 95-325536/42.
DR N-PSDB; Q94347.
PT Interleukin-6 mutant, related DNA and expression vectors - has
PT higher stability than natural interleukin-6
PS Claim 1; Pages 14-15; 18pp; Japanese.
CC Q94347 encodes R77392 the human IL-6 deletion mutant IL-6a/C2
CC which lacks the amino acids Gly5-Leu19 and Cys73-Cys83 of the
CC wild type protein. The cDNA can be used for the recombinant prodn.
CC of IL-6a/C2 which has increased stability compared to wild type
CC IL-6.

```

SQ Sequence 158 AA;

Query Match 64.7%; Score 44; DB 15; Length 158;  
Best Local Similarity 55.6%; Pred. No. 2.44e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 17 ilddgisalr 25

:|||:||

QY 1 LLDGTATLR 9

Search completed: Tue Jun 10 11:22:16 1997  
Job time : 10 secs.

\*\*\*\*\*  
MPSRCH\_PP  
(TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:21:36 1997; MasPar time 2.75 Seconds  
Tabular output not generated. 103.732 Million cell updates/sec

Title: >US-08-231-565A-31  
Description: (1-10) from US08231565A.pep  
Perfect Score: 68  
Sequence: 1 LLDGTATLRL 10

Scoring table: PAM 150  
Gap 15

Searched: 89912 seqs, 28507787 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p1r50

1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 23.639; Variance 32.253; scale 0.733

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	68	100.0	661	13	glycoprotein gp100 p	2.96e-03
2	68	100.0	668	13	melanocyte-specific	2.96e-03
3	51	75.0	811	13	connectin 1 - chick	8.16e+00
4	50	73.5	431	9	dioxigenase beta cha	1.25e+01
5	50	73.5	766	11	RAS GTPase-activatin	1.25e+01
6	50	73.5	846	7	neurofilament triple	1.25e+01
7	50	73.5	858	7	neurofilament triple	1.25e+01
8	49	72.1	184	7	fibrin type 1 - Sal	1.89e+01
9	49	72.1	185	7	type 1 fibrin prot	1.89e+01
10	49	72.1	376	9	crtf protein - Rhodo	1.89e+01
11	48	70.6	34	11	ascorbic acid-induci	2.85e+01
12	48	70.6	240	11	cell division contro	2.85e+01
13	48	70.6	240	11	cdcl4 protein - firs	2.85e+01
14	48	70.6	318	10	deacetylcephalospori	2.85e+01
15	48	70.6	433	16	aspartate aminotrans	2.85e+01
16	48	70.6	433	5	aspartate transamina	2.85e+01
17	48	70.6	453	16	aspartate aminotrans	2.85e+01
18	48	70.6	453	1	aspartate transamina	2.85e+01
19	48	70.6	454	1	aspartate transamina	2.85e+01
20	48	70.6	455	5	aspartate aminotrans	2.85e+01
21	48	70.6	463	16	aspartate aminotrans	2.85e+01

22	48	70.6	463	16	539926	aspartate aminotrans	2.85e+01
23	48	70.6	463	5	S33528	aspartate transamina	2.85e+01
24	47	69.1	54	11	S56725	ascorbic acid-induci	4.28e+01
25	47	69.1	190	7	S12674	ribosomal protein S9	4.28e+01
26	47	69.1	351	11	S56716	protein kinase 3 - s	4.28e+01
27	47	69.1	354	11	S60611	probable serine/thre	4.28e+01
28	47	69.1	359	11	S56720	probable serine/thre	4.28e+01
29	47	69.1	362	11	S56718	protein kinase 1 - A	4.28e+01
30	47	69.1	363	11	S36944	probable serine/thre	4.28e+01
31	47	69.1	1402	16	S62557	hypothetical protein	4.28e+01
32	46	67.6	318	9	C64066	transcriptional regu	6.37e+01
33	46	67.6	332	11	A46408	ascorbic acid-induci	6.37e+01
34	46	67.6	345	14	I49113	vascular cell adhesi	6.37e+01
35	46	67.6	345	14	A46052	vascular cell adhesi	6.37e+01
36	46	67.6	382	8	S12355	xcp protein - pseud	6.37e+01
37	46	67.6	538	14	JC2457	vascular cell adhesi	6.37e+01
38	46	67.6	739	14	A48919	vascular cell adhesi	6.37e+01
39	46	67.6	739	14	JN0581	vascular cell adhesi	6.37e+01
40	46	67.6	739	14	J50675	vascular cell adhesi	6.37e+01
41	46	67.6	818	5	JC4058	fibroblast growth fa	6.37e+01
42	46	67.6	822	5	B49151	fibroblast growth fa	6.37e+01
43	46	67.6	822	5	S19947	fibroblast growth fa	6.37e+01
44	46	67.6	3412	4	GNWV7B	genome polyprotein -	6.37e+01
45	46	67.6	3414	4	GNWVNE	genome polyprotein -	6.37e+01

ALIGNMENTS

RESULT ENTRY	1	A53668	#type complete
TITLE		glycoprotein gp100 precursor, melanocyte lineage - human	
ALTERNATE_NAMES		melanoma antigen 25	
ORGANISM		#formal_name Homo sapiens #common_name man	
DATE		07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 25-May-1996	
ACCESSIONS		A53668; A55753	
REFERENCE		A53668	
#authors		Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G.	
#journal		J. Biol. Chem. (1994) 269:20126-20133	
#title		Molecular characterization of the melanocyte lineage-specific antigen gp100.	
#accession		A53668	
#molecule_type		mRNA	
#residues		1-661	#label ADE
REFERENCE		A55753	
#authors		Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Yannelli, J.R.; Adema, G.J.; Miki, T.; Rosenberg, S.A.	
#journal		Proc. Natl. Acad. Sci. U.S.A. (1994) 91:6458-6462	
#title		Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection.	
#accession		A55753	
#status		nucleic acid sequence not shown; not compared with conceptual translation	
#molecule_type		mRNA	
#residues		1-161, F', 163-661	#label KAW
KEYWORDS		glycoprotein	
SUMMARY		#length 661 #molecular-weight 70255 #checksum 5487	
Query Match		100.08; Score 68; DB 13; Length 661;	
Best Local Similarity		100.08; Pred. No. 2.96e-03;	
Matches		10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	457	lldgtatlrl 466	
QY	1	LLDGTATLRL 10	
RESULT ENTRY	2	A41234	#type complete
TITLE		melanocyte-specific protein Pmel-17 precursor - human	

```

ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
              30-Sep-1993
ACCESSIONS    A41234
REFERENCE      A41234
#authors      Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
              Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
              Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
              coat color locus on mouse chromosome 10 and is in a
              syntenic region on human chromosome 12.
#cross-references MUID:92021023
#accession    A41234
#status       preliminary
#molecule_type mRNA
#residues     1-668 ##label KWO
#cross-references GB:M7348
SUMMARY       #length 668 #molecular-weight 70932 #checksum 6409
Query Match   100.0%; Score 68; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.96e-03;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 457 lldgtatlrl 466
|||||
Qy 1 LLDGTATLRL 10

RESULT 3
ENTRY   #type fragment
TITLE   connectin 1 - Chicken (fragment)
ORGANISM #formal_name Callus gallus #common_name chicken
DATE    14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change
03-Mar-1995
ACCESSIONS PN0689
REFERENCE   PN0689
#authors    Maruyama, K.; Endo, T.; Kume, H.; Kawamura, Y.; Kanzawa, N.;
              Kimura, S.; Kawashima, S.; Maruyama, K.
#journal     J. Biochem. (1994) 115:147-149
#title       A partial connectin cDNA encoding a novel type of RSP motifs
              isolated from chicken embryonic skeletal muscle.
#accession   PN0689
#molecule_type mRNA
#residues    1-811 ##label MAR
#experimental_source embryonic skeletal muscle
COMMENT      This protein is a gigantic elastic protein.
SUMMARY       #length 811 #checksum 9464
Query Match   75.0%; Score 51; DB 13; Length 811;
Best Local Similarity 70.0%; Pred. No. 8.16e-00;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 101 lqdgasrl 110
|||||
Qy 1 LLDGTATLRL 10

RESULT 4
ENTRY   #type complete
TITLE   dioxygenase beta chain homolog - Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE    16-Feb-1995 #sequence_revision 12-May-1995 #text_change
23-Feb-1996
ACCESSIONS S49292
REFERENCE   S49292
#authors    Turin, E.; Gasser, F.; Biville, F.
#submission submitted to the EMBL Data Library, September 1994
#accession   S49292
#title       Cloning and sequencing of an E. coli gene homologous to
              dioxygenase of Gram negative bacteria.
#accession   S49292
#molecule_type DNA
#residues    1-431 ##label TUR

```

```

#cross-references EMBL:237966
SUMMARY       #length 431 #molecular-weight 48138 #checksum 9201
Query Match   73.5%; Score 50; DB 9; Length 431;
Best Local Similarity 77.8%; Pred. No. 1.25e-01;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 316 lngtatlr 324
|||||
Qy 2 LDGTATLRL 10

RESULT 5
ENTRY   #type complete
TITLE   RAS GTPase-activating protein sar1 - fission yeast
              (Schizosaccharomyces pombe)
ORGANISM #formal_name Schizosaccharomyces pombe
DATE    06-Dec-1991 #sequence_revision 06-Dec-1991 #text_change
08-Dec-1995
ACCESSIONS A40258; A56606
REFERENCE   A40258
#authors    Imai, Y.; Miyake, S.; Hughes, D.A.; Yamamoto, M.
#journal     Mol. Cell. Biol. (1991) 11:3088-3094
#title       Identification of a GTPase-activating protein homolog in
              Schizosaccharomyces pombe.
#cross-references MUID:91246176
#accession   A40258
#status       preliminary
#molecule_type DNA
#residues    1-766 ##label IMA
REFERENCE   A56606
#authors    Wang, Y.; Boguski, M.; Riggs, M.; Rodgers, L.; Wigler, M.
#journal     Cell Regul. (1991) 2:453-465
#title       sar1, a gene from Schizosaccharomyces pombe encoding a
              protein that regulates ras1.
#cross-references MUID:91355280
#accession   A56606
#status       preliminary
#molecule_type DNA
#residues    1-766 ##label WAN
#cross-references NCBI:106678; NCBI:106679
#note        sequence extracted from NCBI backbone
CLASSIFICATION #superfamily ras-specific GAP catalytic domain homology
FEATURE         154-398
SUMMARY       #domain ras-specific GAP catalytic domain homology
              #label GAP
              #length 766 #molecular-weight 87539 #checksum 1803
Query Match   73.5%; Score 50; DB 11; Length 766;
Best Local Similarity 70.0%; Pred. No. 1.25e-01;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 565 lldelstlrl 574
|||||
Qy 1 LLDGTATLRL 10

RESULT 6
ENTRY   #type complete
TITLE   neurofilament triplet M protein - chicken
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE    07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
18-Jun-1993
ACCESSIONS S08061
REFERENCE   S08061
#authors    Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.
#submission submitted to the EMBL Data Library, November 1989
#accession   S08061
#molecule_type DNA
#residues    1-846 ##label ZOP
#cross-references EMBL:X17102
#introns     355/3; 385/2

```

```

CLASSIFICATION #superfamily cytoskeletal keratin
KEYWORDS       coiled coil
SUMMARY        #length 846 #molecular-weight 94428 #checksum 400

Query Match    73.5%; Score 50; DB 7; Length 846;
Best Local Similarity 60.0%; Pred. No. 1.25e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 82 llnaaaelkl 91
   ||:|:| ||:|
QY 1 LLDGTATLRL 10

RESULT 7
ENTRY #S15762 #type complete
TITLE neurofilament triplet M protein - chicken
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
12-Apr-1995
ACCESSIONS S15762
REFERENCE S15762
#authors Zopf, D.; Dineva, B.; Betz, H.; Gundelfinger, E.D.
#journal Nucleic Acids Res. (1990) 18:521-529
#title Isolation of the chicken middle-molecular weight
        neurofilament (NF-M) gene and characterization of its
        promoter.
#cross-references MUID:90174973
#accession S15762
#status preliminary
#molecule_type DNA
#residues 1-858 #label ZOP
#cross-references EMBL:x17102
GENETICS
#introns 355/3; 397/2
CLASSIFICATION #superfamily cytoskeletal keratin
SUMMARY        #length 858 #molecular-weight 95834 #checksum 8117

Query Match    73.5%; Score 50; DB 7; Length 858;
Best Local Similarity 60.0%; Pred. No. 1.25e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 82 llnaaaelkl 91
   ||:|:| ||:|
QY 1 LLDGTATLRL 10

RESULT 8
ENTRY #S20682 #type complete
TITLE fimbrin type 1 - Salmonella typhi
ORGANISM #formal_name Salmonella typhi
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
12-Apr-1995
ACCESSIONS S20682
REFERENCE S20682
#authors Rossolini, G.M.; Muscas, P.; Chiesurin, A.; Satta, G.
#submission Submitted to the EMBL Data Library, March 1992
#accession S20682
#status preliminary
#molecule_type DNA
#residues 1-184 #label ROS
#cross-references EMBL:x65168
CLASSIFICATION #superfamily type 1 fimbrin protein
SUMMARY        #length 184 #molecular-weight 18793 #checksum 3117

Query Match    72.1%; Score 49; DB 7; Length 184;
Best Local Similarity 60.0%; Pred. No. 1.89e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 149 lvegtntlrf 158
   ||:|:| ||:|
QY 1 LLDGTATLRL 10

CLASSIFICATION #superfamily cytoskeletal keratin
SUMMARY        #length 846 #molecular-weight 94428 #checksum 400

Query Match    73.5%; Score 50; DB 7; Length 846;
Best Local Similarity 60.0%; Pred. No. 1.25e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 82 llnaaaelkl 91
   ||:|:| ||:|
QY 1 LLDGTATLRL 10

RESULT 9
ENTRY #B28393 #type complete
TITLE type 1 fimbrin protein precursor - Salmonella typhimurium
ORGANISM #formal_name Salmonella typhimurium
DATE 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change
19-Oct-1995
ACCESSIONS B28393; A05121
REFERENCE A91858
#authors Purcell, B.K.; Pruckler, J.; Clegg, S.
#journal J. Bacteriol. (1987) 169:5831-5834
#title Nucleotide sequences of the genes encoding type 1 fimbrin
        subunits of Klebsiella pneumoniae and Salmonella
        typhimurium.
#cross-references MUID:88058806
#accession B28393
#molecule_type DNA
#residues 1-185 #label PUR
#note the authors translated the codon GCC for residue 18 as
        Gly, GTG for residue 31 as Ser, ACC for residue 93 as
        Asn, and ACC for residue 107 as Asn
REFERENCE A05121
#authors Waalen, K.; Sletten, K.; Froholm, L.O.; Vaisanen, V.;
        Korhonen, T.K.
#journal FEBS Microbiol. Lett. (1983) 16:149-151
#accession A05121
#molecule_type protein
#residues 23-30,'S',32-45,'X',47-50 #label WAA
CLASSIFICATION #superfamily type 1 fimbrin protein
FEATURE        1-22 #domain signal sequence #label SIG\
23-185 #product type 1 fimbrin protein #label MAT
SUMMARY        #length 185 #molecular-weight 18867 #checksum 4582

Query Match    72.1%; Score 49; DB 7; Length 185;
Best Local Similarity 60.0%; Pred. No. 1.89e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 149 lvegtntlrf 158
   ||:|:| ||:|
QY 1 LLDGTATLRL 10

RESULT 10
ENTRY #S49626 #type complete
TITLE ctfF protein - Rhodobacter sphaeroides
ORGANISM #formal_name Rhodobacter sphaeroides
DATE 05-Mar-1995 #sequence_revision 12-May-1995 #text_change
12-May-1995
ACCESSIONS S49626
REFERENCE S49619
#authors Lang, H.P.; Cogdell, R.J.; Takaichi, S.; Hunter, C.N.
#submission submitted to the EMBL Data Library, November 1994
#description The complete DNA sequence, specific TNS insertion map and
        gene assignment of the carotenoid biosynthesis genes of
        Rhodobacter sphaeroides.
#accession S49626
#status preliminary
#molecule_type DNA
#residues 1-376 #label LAN
#cross-references EMBL:X82458
GENETICS
#gene ctfF
SUMMARY        #length 376 #molecular-weight 40227 #checksum 9160

Query Match    72.1%; Score 49; DB 9; Length 376;
Best Local Similarity 60.0%; Pred. No. 1.89e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 94 lllqgaalxl 103
   ||:|:| ||:|
QY 1 LLDGTATLRL 10

```

```

RESULT 11
ENTRY S56724 #type fragment
TITLE abscisic acid-inducible protein kinase homolog (clone AspK4)
ORGANISM - oat (fragment)
DATE #formal_name Avena sativa #common_name oat
27-Oct-1995 #sequence_revision 19-Jan-1996 #text_change
19-Jan-1996
ACCESSIONS S56724
REFERENCE S56638
#authors Huttly, A.K.; Phillips, A.L.
#journal Plant Mol. Biol. (1995) 27:1043-1052
#title Gibberellin-regulated expression in oat aleurone cells of two
kinases that show homology to MAP kinase and a ribosomal
protein kinase.
#accession S56724
##molecule_type mRNA
##residues 1-54 ##label HUT
KEYWORDS phosphotransferase; protein kinase
SUMMARY #length 54 #checksum 3285
Query Match 70.6%; Score 48; DB 11; Length 54;
Best Local Similarity 60.0%; Pred. No. 2.85e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 3 lldgstapl 12
|:::|
QY 1 LLDGTATLRL 10

RESULT 12
ENTRY A47731 #type complete
TITLE cell division control protein cdc14 - fission yeast
ALTERNATE_NAMES p28(cdc14)
ORGANISM (Schizosaccharomyces pombe)
DATE #formal_name Schizosaccharomyces pombe
27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change
27-Jun-1994
ACCESSIONS A47731
REFERENCE A47731
#authors Fankhauser, C.; Simanis, V.
#journal Mol. Biol. Cell (1993) 4:531-539
#title The Schizosaccharomyces pombe cdc14 gene is required for
septum formation and can also inhibit nuclear division.
#accession A47731
##status preliminary
##molecule_type DNA
##residues 1-240 ##label FAN
##cross-references EMBL:X72911
GENETICS cdc14
#gene cell division control; mitosis
KEYWORDS #length 240 #molecular-weight 28160 #checksum 9965
SUMMARY
Query Match 70.6%; Score 48; DB 11; Length 240;
Best Local Similarity 60.0%; Pred. No. 2.85e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 111 lfedkatlrl 120
|:::|
QY 1 LLDGTATLRL 10

RESULT 13
ENTRY S35794 #type complete
TITLE cdc14 protein - fission yeast (Schizosaccharomyces pombe)
ORGANISM #formal_name Schizosaccharomyces pombe
DATE #formal_name Schizosaccharomyces pombe
13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
13-Jan-1995
ACCESSIONS S35794
REFERENCE S35794
#authors Simanis, V.
#submission submitted to the EMBL Data Library, March 1993
#accession S35794

```

```

##status preliminary
##molecule_type DNA
##residues 1-240 ##label SIM
##cross-references EMBL:X72911
SUMMARY #length 240 #molecular-weight 28160 #checksum 9965
Query Match 70.6%; Score 48; DB 11; Length 240;
Best Local Similarity 60.0%; Pred. No. 2.85e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 111 lfedkatlrl 120
|:::|
QY 1 LLDGTATLRL 10

RESULT 14
ENTRY A39204 #type complete
TITLE deacetylcephalosporin hydroxylase - Streptomyces clavuligerus
ORGANISM #formal_name Streptomyces clavuligerus
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
22-Jul-1994
ACCESSIONS A39204; A23713
REFERENCE A39204
#authors Kovacevic, S.; Miller, J.R.
#journal J. Bacteriol. (1991) 173:398-400
#title Cloning and sequencing of the beta-lactam hydroxylase gene
(ceff) from Streptomyces clavuligerus: gene duplication may
have led to separate hydroxylase and expandase activities
in the actinomycetes.
#cross-references MUID:91100311
#accession A39204
##molecule_type DNA
##residues 1-318 ##label KOV
##cross-references GB:M37186
REFERENCE A23713
#authors Baker, B.J.; Dotzlauf, J.E.; Yeh, W.K.
#journal J. Biol. Chem. (1991) 266:5087-5093
#title Deacetylcephalosporin C hydroxylase of Streptomyces
clavuligerus. Purification, characterization,
bifunctionality, and evolutionary implication.
#cross-references MUID:91161600
#accession A23713
##molecule_type protein
##residues 2-29;92-100 ##label BAK
GENETICS cefF
#gene
SUMMARY #length 318 #molecular-weight 34584 #checksum 3583
Query Match 70.6%; Score 48; DB 10; Length 318;
Best Local Similarity 60.0%; Pred. No. 2.85e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 156 lldadpvlrl 165
|:::|
QY 1 LLDGTATLRL 10

RESULT 15
ENTRY S39928 #type complete
TITLE aspartate aminotransferase AAT2-c2 - alfalfa
ORGANISM #formal_name Medicago sativa #common_name alfalfa
DATE 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
07-Oct-1994
ACCESSIONS S39928
REFERENCE S39925
#authors Gregerson, R.G.; Petrowski, M.; Larson, R.L.; Gantt, J.S.;
Vance, C.P.
#journal Mol. Gen. Genet. (1993) 241:124-128
#title Molecular analysis of allelic polymorphism at the AAT2 locus
of alfalfa.
#accession S39928
##status preliminary
##residues 1-453 ##label GRE

```

SUMMARY            #length 453   #molecular-weight 49693   #checksum 3422

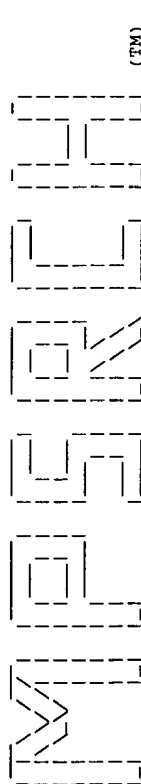
Query Match        70.6%;   Score 48;   DB 16;   Length 453;  
Best Local Similarity 66.7%;   Pred. NO. 2.85e+01;  
Matches        6;   Conservative    2;   Mismatches   1;   Indels    0;   Gaps    0;

Db    154 lsgtgsrlrl 162  
      | ||:||||  
Qy    2 LDGTATLRRL 10

Search completed: Tue Jun 10 11:21:48 1997  
Job time : 12 secs.



\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:21:08 1997; MasPar time 1.95 Seconds  
Tabular output not generated.  
108.824 Million cell updates/sec

Title: >US-08-231-565A-31  
Description: (1-10) from US08231565A.pep  
Perfect Score: 68  
Sequence: 1 LLDGTATIRL 10

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 24.519; Variance 26.804; scale 0.915

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	68	100.0	668	7	PM17_HUMAN PMEL 17 PROTEIN PRECU	1.24e+04
2	56	82.4	796	11	YS8A_CAEEL HYPOTHETICAL 84.3 KD	1.26e+01
3	50	73.5	330	2	CD86_RABIT B LYMPHOCYTE ACTIVATI	2.91e+00
4	50	73.5	766	4	GAP1_SCHPO GTPASE-ACTIVATING PRO	2.91e+00
5	50	73.5	857	6	NFM_CHICK NEUROFILAMENT TRIPLET	2.91e+00
6	49	72.1	184	4	PM1A_SALT1 TYPE-1 FIMBRIAL PROTE	4.79e+00
7	49	72.1	185	4	PM1A_SALT1 FIMBRIAL SUBUNIT TYPE	4.79e+00
8	49	72.1	185	4	PM1A_SALT1 TYPE-1 FIMBRIAL PROTE	4.79e+00
9	49	72.1	376	2	CR1F_RHOSH HYDROXYNEUROSPORINE M	4.79e+00
10	48	70.6	240	2	CC14_SCHPO CELL DIVISION CONTROL	7.82e+00
11	48	70.6	315	11	YX25_MYCTU PROBABLE INTEGRASE/RE	7.82e+00
12	48	70.6	318	2	CEFF_STRCL DEACETOXYCEPHALOSPORI	7.82e+00
13	48	70.6	453	1	AATM_ARATH ASPARTATE AMINOTRANSF	7.82e+00
14	48	70.6	454	1	AATM_LUPAN ASPARTATE AMINOTRANSF	7.82e+00
15	48	70.6	3491	3	ERY1_SACER ERYTHRONOLIDE SYNTHAS	7.82e+00
16	47	69.1	190	8	RS9_TRYBB PROBABLE 40S RIBOSOMA	1.27e+01
17	47	69.1	363	1	ASK1_ARATH SERINE/THREONINE-PROT	1.27e+01
18	47	69.1	485	9	SAHE_PHASS ADENOSYLHOMOCYSTEINAS	1.27e+01
19	47	69.1	497	9	SR52_HORVU SIGNAL RECOGNITION PA	1.27e+01
20	47	69.1	497	9	SR51_HORVU SIGNAL RECOGNITION PA	1.27e+01
21	47	69.1	504	10	Y05N_MYCTU HYPOTHETICAL 53.5 KD	1.27e+01
22	47	69.1	1402	1	ATCX_SCHPO PROBABLE CALCIUM-TRAN	1.27e+01

23	46	67.6	21	6	LPRM_CORDI	23S RRNA METHYLASE LE	2.03e+01
24	46	67.6	318	10	TYRR_HAEIN	TRANSCRIPTIONAL REGUL	2.03e+01
25	46	67.6	332	1	AAIP_WHEAT	ABCSIC ACID-INDUCIBL	2.03e+01
26	46	67.6	343	11	YX36_MYCTU	HYPOTHETICAL 36.7 KD	2.03e+01
27	46	67.6	382	4	GSPL_PSEAE	GENERAL SECRETION PAT	2.03e+01
28	46	67.6	570	11	YNE3_CAEEL	HYPOTHETICAL 54.2 KD	2.03e+01
29	46	67.6	719	11	YMP8_CAEEL	HYPOTHETICAL 82.6 KD	2.03e+01
30	46	67.6	739	10	VCA1_RAT	VASCULAR CELL ADHESIO	2.03e+01
31	46	67.6	739	10	VCA1_MOUSE	VASCULAR CELL ADHESIO	2.03e+01
32	46	67.6	997	10	VGNM_APMV	GENOME POLYPROTEIN M	2.03e+01
33	46	67.6	3412	7	POLG_TBEVS	GENOME POLYPROTEIN (C	2.03e+01
34	46	67.6	3414	7	POLG_TBEVW	GENOME POLYPROTEIN (C	2.03e+01
35	46	67.6	3414	7	POLG_TBEVH	GENOME POLYPROTEIN (C	2.03e+01
36	45	66.2	220	2	CMBB_ECOLI	HEME EXPORTER PROTEIN	3.23e+01
37	45	66.2	231	2	CASB_MOUSE	BETA CASEIN PRECURSOR	3.23e+01
38	45	66.2	232	4	GUS8_RAT	POSSIBLE GUSTATORY RE	3.23e+01
39	45	66.2	310	1	ADPR_LACIA	ATP-DEPENDENT PROTEAS	3.23e+01
40	45	66.2	522	10	UNC7_CAEEL	UNC-7 PROTEIN.	3.23e+01
41	45	66.2	633	7	PABP_XENLA	POLYADENYLATE-BINDING	3.23e+01
42	45	66.2	828	11	YF44_YEAST	HYPOTHETICAL 95.4 KD	3.23e+01
43	45	66.2	882	7	PM11_HUMAN	PROBABLE TRANSCRIPTIO	3.23e+01
44	45	66.2	1271	1	BCR_HUMAN	BREAKPOINT CLUSTER RE	3.23e+01
45	45	66.2	2554	1	7LES_DROME	SEVENLESS PROTEIN (EC	3.23e+01

ALIGNMENTS

RESULT	1	PM17_HUMAN	STANDARD;	PRT;	668 AA.
AC	P40967;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92021023.				
RA	KWON B.S., CHINTAMANANI C., KOZAK C.A., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N., BARTON D., FRANKCE U., KOBAYASHI Y.,				
RA	KIM K.-K.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).				
CC	-1- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	-1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; G190106; -.				
DR	MM: 155550; -.				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23	POTENTIAL.	
FT	CHAIN	24	668	PMEL 17 PROTEIN.	
FT	TRANSMEM	575	595	POTENTIAL.	
FT	TRANSMEM	603	623	POTENTIAL.	
FT	DOMAIN	217	307	PKD.	
FT	DOMAIN	315	444	10 X 13 AA TANDEM REPEATS.	
FT	REPEAT	315	327	1.	
FT	REPEAT	328	340	2.	
FT	REPEAT	341	353	3.	
FT	REPEAT	354	366	4.	
FT	REPEAT	367	379	5.	
FT	REPEAT	390	392	6.	
FT	REPEAT	393	405	7.	
FT	REPEAT	406	418	8.	
FT	REPEAT	419	431	9.	
FT	REPEAT	432	444	10.	
FT	CARBOHYD	81	81	POTENTIAL.	
FT	CARBOHYD	106	106	POTENTIAL.	
FT	CARBOHYD	111	111	POTENTIAL.	
FT	CARBOHYD	321	321	POTENTIAL.	
FT	CARBOHYD	568	568	POTENTIAL.	
FT	SEQUENCE	668 AA;	70992 MW;	6E8E1AFO CRC32;	

Query Match 100.0%; Score 68; DB 7; Length 668;  
 Best Local Similarity 100.0%; Pred. No. 1.24e-04;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 457 lldgtatrl 466  
 Qy 1 LLDGTATLRL 10

RESULT 2  
 ID YS8A\_CAEEL STANDARD; PRT; 796 AA.

AC Q09625;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 33, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 84.3 KD PROTEIN ZK945.10 IN CHROMOSOME II.  
 GN ZK945.10.

OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA WILKINSON-SPROAT J.;  
 RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; Z48582; G695475; -;  
 DR EMBL; Z48544; G695502; -;

DR WORMPEP; ZK945.10; CE01732.  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
 FT TRANSMEM 11 30 POTENTIAL.  
 FT DOMAIN 273 546 SER/THR-RICH.  
 FT DOMAIN 656 752 SER/THR-RICH.

FT SEQUENCE 796 AA; 84306 MW; 8099740C CRC32;  
 SQ SEQUENCE 796 AA; 84306 MW; 8099740C CRC32;

Query Match 82.4%; Score 56; DB 11; Length 796;  
 Best Local Similarity 80.0%; Pred. No. 1.26e-01;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 72 lldgtatrl 81  
 Qy 1 LLDGTATLRL 10

RESULT 3  
 ID CD86\_RABIT STANDARD; PRT; 330 AA.

AC P42071;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE B LYMPHOCYTE ACTIVATION ANTIGEN CD86 PRECURSOR (ACTIVATION B7-2 ANTIGEN).

GN CD86.  
 OS ORYCTOLAGUS CUNICULUS (RABBIT).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; LAGOMORPHA.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-B/J X CHB:HM;  
 RX MEDLINE; 95369849.

RA ISONO T.; SETO A.;  
 RL IMMUNOGENETICS 42:217-220(1995).

CC -!- FUNCTION: RECEPTOR INVOLVED IN THE COSTIMULATORY SIGNAL ESSENTIAL FOR T LYMPHOCYTE PROLIFERATION AND INTERLEUKIN 2 PRODUCTION, BY BINDINGS CD28 OR CTLA-4. MAY PLAY A CRITICAL ROLE IN THE EARLY EVENTS OF T CELL ACTIVATION AND COSTIMULATION OF NAIVE T CELLS, SUCH AS DECIDING BETWEEN IMMUNITY AND ANERGY THAT IS MADE BY T CELLS WITHIN 24 HOURS AFTER ACTIVATION.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS ONE C2-LIKE AND ONE V-LIKE DOMAIN.

DR EMBL; D49842; G755099; -;

KW IMMUNOGLOBULIN FOLD; T-CELL; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE; RECEPTOR.

FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 330  
 FT DOMAIN 23 247 B LYMPHOCYTE ACTIVATION ANTIGEN CD86.  
 FT TRANSMEM 248 268 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 269 330 POTENTIAL.  
 FT DOMAIN 33 117 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 149 225 IG-LIKE V-TYPE DOMAIN.  
 FT DISULFID 40 110 IG-LIKE C2-TYPE DOMAIN.  
 FT DISULFID 157 218 POTENTIAL.  
 FT CARBOHYD 33 33 POTENTIAL.  
 FT CARBOHYD 135 135 POTENTIAL.  
 FT CARBOHYD 146 146 POTENTIAL.  
 FT CARBOHYD 154 154 POTENTIAL.  
 FT CARBOHYD 177 177 POTENTIAL.  
 FT CARBOHYD 192 192 POTENTIAL.  
 FT CARBOHYD 198 198 POTENTIAL.  
 FT CARBOHYD 213 213 POTENTIAL.  
 SQ SEQUENCE 330 AA; 37142 MW; 9A3CD9C8 CRC32;

Query Match 73.5%; Score 50; DB 2; Length 330;  
 Best Local Similarity 60.0%; Pred. No. 2.91e+00;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 19 llsagaaslr 28  
 Qy 1 LLDGTATLRL 10

RESULT 4

ID GAP1\_SCHPO STANDARD; PRT; 766 AA.

AC P33277;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GTPASE-ACTIVATING PROTEIN.

GN GAP1 OR SRC1 OR SARI.  
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE; 91246176.  
 RA IMAI Y.; MIYAKE S.; HUGHES D.A.; YAMAMOTO M.;

RL MOL. CELL. BIOL. 11:3088-3094(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RX MEDLINE; 91355280.  
 RA WANG Y.; BOGUSKI M.; RIGGS M.; RODGERS L.; WIGLER M.;

RL CELL REGUL. 2:453-465(1991).  
 CC -!- FUNCTION: INHIBITORY REGULATOR OF THE RAS-CYCCLIC AMP

CC -!- PATHWAY IN S.POMBE. STIMULATES THE GTPASE ACTIVITY OF RASL.

CC -!- SIMILARITY: TO OTHER RAS GTPASE-ACTIVATING PROTEINS.

DR EMBL; D10457; G218535; -;

DR EMBL; S37449; G234783; -;

DR PIR; A40258; A40258.  
 DR PROSITE; PS00509; RAS\_GTPASE\_ACTIV\_1.

DR PROSITE; PS50018; RAS\_GTPASE\_ACTIV\_2.  
 KW GTPASE ACTIVATION.

FT DOMAIN 167 379 RAS-GAP.  
 SQ SEQUENCE 766 AA; 87540 MW; E2BA368D CRC32;

Query Match 73.5%; Score 50; DB 4; Length 766;  
 Best Local Similarity 70.0%; Pred. No. 2.91e+00;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 565 lldelstlrl 574  
 Qy 1 LLDGTATLRL 10

RESULT 5

ID NFM\_CHICK STANDARD; PRT; 857 AA.

AC P16053;  
 DT 01-APR-1990 (REL. 14, CREATED)

DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)

01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE NEUROFILAMENT TRIPLET M PROTEIN (160 KD NEUROFILAMENT PROTEIN) (NF-M).  
 GN NEFM.  
 OS GALLUS GALLUS (CHICKEN).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;  
 CC GALLIFORMES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90174973.  
 RA ZOPF D., DINEVA B., BETZ H., GUNDELFINGER E.D.;  
 RL NUCLEIC ACIDS RES. 18:521-529(1990).  
 RN [2]  
 RP SEQUENCE OF 259-857 FROM N.A.  
 RX MEDLINE; 88112814.  
 RA ZOPF D., HERMANS-BORGMEYER I., GUNDELFINGER E.D., BETZ H.;  
 RL GENES DEV. 1:699-708(1987).  
 CC -!- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,  
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.  
 CC -!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPEPTIDE K-S-P. NFM IS  
 CC PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS  
 CC THOUGHT THAT PHOSPHORYLATION OF NFM RESULTS IN THE FORMATION OF  
 CC INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE  
 CC OF AXONAL CALIBER.  
 CC -!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING  
 CC OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE  
 CC LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND  
 CC COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.  
 CC -!- SIMILARITY: TO ALL OTHER INTERMEDIATE FILAMENT PROTEINS.  
 DR EMBL; X17102; G63689; -.  
 DR EMBL; X05558; G63686; -.  
 DR PIR; A27040; A27040.  
 DR PIR; S08061; S08061.  
 DR PIR; S15762; S15762.  
 DR PROSITE; PS00226; IF.  
 KW INTERMEDIATE FILAMENT; HEPTAD REPEAT PATTERN; COILED COIL; NEURONE;  
 FT INIT\_MET 0 0  
 FT DOMAIN 1 98 HEAD.  
 FT DOMAIN 99 406 ROD.  
 FT DOMAIN 407 857 TAIL.  
 FT DOMAIN 99 130 COIL 1A.  
 FT DOMAIN 131 143 LINKER 1.  
 FT DOMAIN 144 242 COIL 1B.  
 FT DOMAIN 243 259 LINKER 12.  
 FT DOMAIN 260 281 COIL 2A.  
 FT DOMAIN 282 285 LINKER 2.  
 FT DOMAIN 286 406 COIL 2B.  
 FT CARBOHYD 46 46 GLCNAC (BY SIMILARITY).  
 FT CARBOHYD 426 426 GLCNAC (BY SIMILARITY).  
 FT CONFLICT 546 546 G -> R (IN REF. 2).  
 SQ SEQUENCE 857 AA; 95704 MW; 3D05FFDD CRC32;  
 Query Match 73.5%; Score 50; DB 6; Length 857;  
 Best Local Similarity 60.0%; Pred. No. 2.91e+00;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 81 llncgaaelkl 90  
 QY 1 LLDGTATLRL 10  
 RESULT 6  
 ID FMIA\_SALTI STANDARD; PRT; 184 AA.  
 AC P37920;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE TYPE-1 FIMBRIAL PROTEIN, A CHAIN PRECURSOR (TYPE-1A PILIN).  
 GN FIMA.  
 OS SALMONELLA TYPHI.  
 CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 CC ENTEROBACTERIACEAE.  
 RN [1]

RC SEQUENCE FROM N.A.  
 RC STRAIN-STV4;  
 RX MEDLINE; 94116831.  
 RA ROSSOLINI G.M., MUSCAS P., CHIESURIN A., SATTA G.;  
 RL FEMS MICROBIOL. LETT. 114:259-266(1993).  
 CC -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING  
 CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5  
 CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO  
 CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.  
 CC -!- SUBCELLULAR LOCATION: FIMBRIA.  
 CC -!- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.  
 DR EMBL; X65168; G47667; -.  
 DR PIR; S20682; S20682.  
 KW FIMBRIA; SIGNAL.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 184 TYPE-1 FIMBRIAL PROTEIN, A CHAIN.  
 FT DISULFID 46 86 PROBABLE.  
 SQ SEQUENCE 184 AA; 18793 MW; 7ED7F4E7 CRC32;  
 Query Match 72.1%; Score 49; DB 4; Length 184;  
 Best Local Similarity 60.0%; Pred. No. 4.79e+00;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 149 lvegtntlrf 158  
 QY 1 LLDGTATLRL 10  
 RESULT 7  
 ID FMIA\_SALTI STANDARD; PRT; 185 AA.  
 AC P53223;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE FIMBRIAL SUBUNIT TYPE 1 PRECURSOR.  
 OS SALMONELLA TYPHIMURIUM.  
 CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 CC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88058806.  
 RA FORCELL B.K., PRUCKLER J., CLEGG S.;  
 RL J. BACTERIOL. 169:5831-5834(1987).  
 CC -!- SIMILARITY: WITH E.COLI AND K.PNEUMONIAE FIMBRIAL SUBUNITS TYPE 1  
 CC PRECURSORS.  
 CC -!- SUBCELLULAR LOCATION: FIMBRIA.  
 CC -!- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.  
 DR EMBL; M18283; G153963; -.  
 KW FIMBRIA; SIGNAL.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 185 FIMBRIAL SUBUNIT TYPE 1.  
 FT DISULFID 46 86 PROBABLE.  
 SQ SEQUENCE 185 AA; 18897 MW; 20D111F7 CRC32;  
 Query Match 72.1%; Score 49; DB 4; Length 185;  
 Best Local Similarity 60.0%; Pred. No. 4.79e+00;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 149 lvegtntlrf 158  
 QY 1 LLDGTATLRL 10  
 RESULT 8  
 ID FMIA\_SALTI STANDARD; PRT; 185 AA.  
 AC P37921;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE TYPE-1 FIMBRIAL PROTEIN, A CHAIN PRECURSOR (TYPE-1A PILIN).  
 GN FIMA.  
 OS SALMONELLA TYPHIMURIUM.  
 CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 CC ENTEROBACTERIACEAE.  
 RN [1]

```

OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RA SWENSON D.L., CLEGG S.;
RL SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 23-50.
RC STRAIN-LT2 / SH6749;
RA WALEN K., SLETTEN K., FROHOLM L.O., VAISANEN V., KORHONEN T.K.;
RL FEMS MICROBIOL. LETT. 16:149-151(1983).
CC -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENT'S RADIATING
CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC -!- SUBCELLULAR LOCATION: FIMBRIA.
CC -!- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.
DR EMBL; L19338; G349131; -
DR STYGENE; SG10275; FIMA.
KW FIMBRIA; SIGNAL.
FT SIGNAL 1 22
FT CHAIN 23 185 TYPE-1 FIMBRIAL PROTEIN, A CHAIN.
FT DISULFID 46 86 PROBABLE.
SQ SEQUENCE 185 AA; 18897 MW; C27656C9 CRC32;

Query Match 72.1%; Score 49; DB 4; Length 185;
Best Local Similarity 60.0%; Pred. No. 4.79e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 149 lvegtntlrf 158
QY 1 LLDGTATLRL 10

RESULT 9
ID CRTF_RHOSH STANDARD; PRT; 376 AA.
AC P54906;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYDROXYNEUROSPORENE METHYLTRANSFERASE (EC 2.1.1.-) (O-METHYLASE).
GN CRTF.
OS RHODOBACTER SPHAEROIDES (RHODORSEUDOMONAS SPHAEROIDES).
OC PROKARYOTA; GRACILICUTES; ANOXYPHOTOBACTERIA; PURPLE BACTERIA;
OC RHODOSPIRILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCIB 8253;
RX MEDLINE; 95238278.
RA LANG H.P., COGDELL R.J., TAKAICHI S., HUNTER C.N.;
RL J. BACTERIOL. 177:2064-2073(1995).
CC -!- FUNCTION: CONVERTS HYDROXYNEUROSPORENE TO METHOXYNEUROSPORENE
CC OR DEMETHYLSPHEROIDENE TO SPHEROIDENE.
CC -!- PATHWAY: CAROTENOID AND CHLOROPHYLL BIOSYNTHESIS.
DR EMBL; X82458; G575413; -
KW PHOTOSYNTHESIS; CHLOROPHYLL BIOSYNTHESIS; CAROTENOID BIOSYNTHESIS;
KW TRANSFERASE; METHYLTRANSFERASE.
SQ SEQUENCE 376 AA; 40227 MW; 0591A671 CRC32;

Query Match 72.1%; Score 49; DB 2; Length 376;
Best Local Similarity 60.0%; Pred. No. 4.79e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 94 llqgggaalkl 103
QY 1 LLDGTATLRL 10

RESULT 10
ID CC14_SCHPO STANDARD; PRT; 240 AA.
AC P36589;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)

```

```

DE CELL DIVISION CONTROL PROTEIN 14.
GN CDC14.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMICETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE; 93326814.
RA FANKHAUSER C., SIMANIS V.;
RL MOL. BIOL. CELL 4:531-539(1993).
CC -!- FUNCTION: MAY PLAY A ROLE IN BOTH THE INITIATION OF MITOSIS AND
CC SEPTUM FORMATION AND, BY DOING SO, BE PART OF THE MECHANISM THAT
CC COORDINATES THESE TWO CELL-CYCLE EVENTS.
DR EMBL; X72911; G312946; -
DR PIR; S35794; S35794.
DR PIR; A47731; A47731.
KW CELL DIVISION; CELL CYCLE; MITOSIS.
SQ SEQUENCE 240 AA; 28160 MW; BC925951 CRC32;

Query Match 70.6%; Score 48; DB 2; Length 240;
Best Local Similarity 60.0%; Pred. No. 7.82e+00;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 111 lfedkatlrl 120
QY 1 LLDGTATLRL 10

RESULT 11
ID YX25_MYCTU STANDARD; PRT; 315 AA.
AC Q10815;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE PROBABLE INTEGRASE/RECOMBINASE CY274.25C.
GN MTCY274.25C.
OS MYCOBACTERIUM TUBERCULOSIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: BELONGS TO THE "PHAGE" INTEGRASE FAMILY.
DR EMBL; Z74024; E248771; -
KW HYPOTHETICAL PROTEIN; DNA RECOMBINATION; DNA INTEGRATION.
FT ACT_SITE 296 296 PROBABLE TRANSIENT COVALENT LINKAGE TO
FT DURING STRAND CLEAVAGE AND REJOINING
FT (BY SIMILARITY).
SQ SEQUENCE 315 AA; 33792 MW; 0C6D806C CRC32;

Query Match 70.6%; Score 48; DB 11; Length 315;
Best Local Similarity 60.0%; Pred. No. 7.82e+00;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 271 lileggadlrv 280
QY 1 LLDGTATLRL 10

RESULT 12
ID CEFF_STRCL STANDARD; PRT; 318 AA.
AC P42220;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE DEACETOXYCEPHALOSPORIN C HYDROXYLASE (EC 1.14.11.-)
DE (DEACETYLCYCEPHALOSPORIN C SYNTHETASE) (DACS) (BETA-LACTAM HYDROXYLASE).
GN CEFF.
OS STREPTOMYCES CLAVULIGERUS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE: 91100311.
RL KOVACEVIC S., MILLER J.R.;
CC 3'-FUNCTION: HYDROXYLATION OF DESACETOXICEPHALOSPORIN C IN
CC 3'-POSITION TO FORM DEACETYLCEPHALOSPORIN C.
CC 3'-PATHWAY: CEPHALOSPORIN BIOSYNTHESIS.
CC 1'-SIMILARITY: STRONG, TO CEFE.
CC 1'-SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF
CC OXIDOREDUCTASES.
DR EMBL; M63809; G153207; -.
KW ANTIBIOTIC BIOSYNTHESIS; OXIDOREDUCTASE; IRON; VITAMIN C.
SQ SEQUENCE 318 AA; 34584 MW; FE91F990 CRC32;

Query Match 70.6%; Score 48; DB 2; Length 318;
Best Local Similarity 60.0%; Pred. No. 7.82e+00;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 156 lldadpvlrl 165
QY 1 LLDGTATLRL 10

RESULT 13
ID AATM_ARATH STANDARD; PRT; 453 AA.
AC P46248;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE ASPARTATE AMINOTRANSFERASE, MITOCHONDRIAL PRECURSOR (EC 2.6.1.1)
DE (TRANSMINASE A).
GN AATI OR ASP5.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPPARALES; CRUCIFERAE.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF;
RA WILKIE S.E., ROPER J., SMITH A., WARREN M.J.;
RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF; STRAIN-CV. LANDSBERG ERRECTA;
RA WILKIE S.E., LAMBERT R., WARREN M.J.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC 1'-FUNCTION: IMPORTANT FOR THE METABOLISM OF AMINO ACIDS AND KREBS-
CC CYCLE RELATED ORGANIC ACIDS. IN PLANTS, IT IS INVOLVED IN NITROGEN
CC METABOLISM AND IN ASPECTS OF CARBON AND ENERGY METABOLISM.
CC 1'- CATALYTIC ACTIVITY: L-ASPARTATE + 2-OXOGLUTARATE = OXALOACETATE +
CC L-GLUTAMATE.
CC 1'- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC 1'- COFACTOR: PYRIDOXAL PHOSPHATE.
CC 1'- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (POTENTIAL).
CC 1'- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
DR EMBL; X81026; G531555; -.
DR EMBL; X91865; G1017411; -.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1.
KW TRANSFERASE; AMINOTRANSFERASE; PYRIDOXAL PHOSPHATE; MITOCHONDRION;
FT TRANSIT PEPTIDE; MULTIGENE FAMILY.
FT CHAIN 53 453 ASPARTATE AMINOTRANSFERASE.
FT BINDING 298 298 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 453 AA; 49803 MW; 0B0381EA CRC32;

Query Match 70.6%; Score 48; DB 1; Length 453;
Best Local Similarity 66.7%; Pred. No. 7.82e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 152 lsgtgsrlrl 160
QY 2 LDGTATLRL 10

```

```

RESULT 14
ID AATM_LUPAN STANDARD; PRT; 454 AA.
AC P26563;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE ASPARTATE AMINOTRANSFERASE-P2, MITOCHONDRIAL PRECURSOR (EC 2.6.1.1)
DE (TRANSMINASE A) (FRAGMENT).
OS LUPINUS ANGSTIFOLIUS (NARROW-LEAVED BLUE LUPINE).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; FABALES;
OC FABACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. UNIHARVEST; TISSUE=ROOT NODULES;
RX MEDLINE; 92322978.
RA REYNOLDS P.H.S., SMITH L.A., JONES W.T., DICKSON J.M.J., JONES S.J.,
RA ROBERTS K., LIDDANE C.P.;
RL PLANT MOL. BIOL. 19:465-472(1992).
CC 1'-FUNCTION: IMPORTANT FOR THE METABOLISM OF AMINO ACIDS AND KREBS-
CC CYCLE RELATED ORGANIC ACIDS. IN PLANTS, IT IS INVOLVED IN NITROGEN
CC METABOLISM AND IN ASPECTS OF CARBON AND ENERGY METABOLISM.
CC 1'- CATALYTIC ACTIVITY: L-ASPARTATE + 2-OXOGLUTARATE = OXALOACETATE +
CC L-GLUTAMATE.
CC 1'- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC 1'- COFACTOR: PYRIDOXAL PHOSPHATE.
CC 1'- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC 1'- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
DR EMBL; X59761; G19139; -.
DR PIR; S16741; ANILB.
DR PIR; S22465; S22465.
DR HSSP; P00508; IAMA.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1.
KW TRANSFERASE; AMINOTRANSFERASE; PYRIDOXAL PHOSPHATE; MITOCHONDRION;
KW TRANSIT PEPTIDE.
FT NON_TER 1 1 MITOCHONDRION.
FT TRANSIT <1 49 ASPARTATE AMINOTRANSFERASE.
FT CHAIN 50 454 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT BINDING 299 299 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 454 AA; 49915 MW; E6419894 CRC32;

Query Match 70.6%; Score 48; DB 1; Length 454;
Best Local Similarity 66.7%; Pred. No. 7.82e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 153 lsgtgsrlrl 161
QY 2 LDGTATLRL 10

RESULT 15
ID ERYI_SACER STANDARD; PRT; 3491 AA.
AC Q03131;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE ERYTHRONOLIDE SYNTHASE, MODULES 1 AND 2 (EC 2.3.1.94) (ORF 1).
GN ERYA.
OS SACCCHAROPOLYSPORA ERYTHRAEA (STREPTOMYCES ERYTHRAEUS).
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; NOCARDIOFORM.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91220065.
RA DONADIO S., STAVER M.J., MCALPINE J.B., SWANSON S.J., KATZ L.;
RL SCIENCE 252:675-679(1991).
RN [2]
RP SEQUENCE OF 3474-3491 FROM N.A.
RX MEDLINE; 93231529.
RA DONADIO S., STAVER M.J.;
RL GENE 126:147-151(1993).
CC 1'- CATALYTIC ACTIVITY: 6 METHYLMALONYL-COA + PROPIONYL-COA = 7 COA
CC + 6-DEOXYERYTHRONOLIDE B.
CC 1'- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN

```

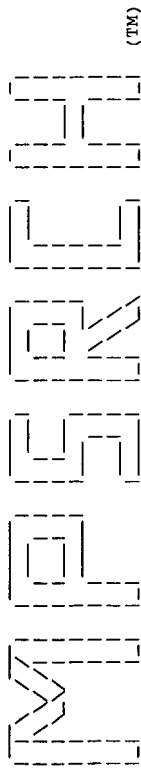
```

CC      BIOSYNTHESIS.
CC      -!- COFACTOR: NADP.
CC      -!- IN EACH OF ERYA TWO COVALENTLY BOUND PHOSPHOPANTHETHEINES.
CC      -!- IN EACH OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A
CC      FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORS CODES
CC      FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE
CC      PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS
CC      REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
CC      AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC      RESPECTIVELY.
CC      -!- BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL
CC      CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR
CC      CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND
CC      ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND
CC      THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-
CC      LENGTH CHAIN.
CC      -!- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
DR      EMBL; M63676; G152692; -.
DR      EMBL; L07626; G294871; -.
DR      PROSITE; PS00112; PHOSPHOPANTHETHEINE.
DR      PROSITE; PS00606; B.KETOACYL SYNTHASE.
KW      TRANSFERASE; ACYLTRANSFERASE; ANTIBIOTIC BIOSYNTHESIS; NADP;
KW      PHOSPHOPANTHETHEINE; MULTIFUNCTIONAL ENZYME.
FT      DOMAIN 1 1972 MODULE 1.
FT      DOMAIN 1979 3491 MODULE 2.
FT      DOMAIN 1 375 ACYLTRANSFERASE (AT).
FT      DOMAIN 419 485 ACYL CARRIER (ACP).
FT      DOMAIN 503 961 BETA-KETOACYL SYNTHASE (KS).
FT      DOMAIN 1030 1356 ACYLTRANSFERASE (AT).
FT      DOMAIN 1611 1794 BETA-KETOACYL REDUCTASE (KR).
FT      DOMAIN 1876 1961 ACYL CARRIER (ACP).
FT      DOMAIN 1979 2441 BETA-KETOACYL SYNTHASE (KS).
FT      DOMAIN 2507 2854 ACYLTRANSFERASE (AT).
FT      DOMAIN 3055 3237 BETA-KETOACYL REDUCTASE (KR).
FT      DOMAIN 3322 3407 ACYL CARRIER (ACP).
FT      ACT_SITE 145 145 ACYL-ENZYME INTERMEDIATE.
FT      BINDING 447 447 PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT      ACT_SITE 677 677 BETA-KETOACYL SYNTHASE.
FT      ACT_SITE 1128 1128 ACYL-ENZYME INTERMEDIATE.
FT      NP_BIND 1614 1660 NADP.
FT      BINDING 1921 1921 PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT      ACT_SITE 2148 2148 BETA-KETOACYL SYNTHASE.
FT      ACT_SITE 2598 2598 ACYL-ENZYME INTERMEDIATE.
FT      NP_BIND 3058 3104 NADP.
FT      BINDING 3367 3367 PHOSPHOPANTHETHEINE (BY SIMILARITY).
SQ      SEQUENCE 3491 AA; 365022 MW; 96D53B98 CRC32;
Query Match 70.6%; Score 48; DB 3; Length 3491;
Best Local Similarity 50.0%; Pred. No. 7.82e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 3376 llnaatglrl 3385
Qy 1 LLDGTATLRL 10

```

Search completed: Tue Jun 10 11:21:17 1997  
Job time : 9 secs.

\*\*\*\*\*


 MPSRCH\_PP (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:23:28 1997; MasPar time 1.91 Seconds

Tabular output not generated. 56.981 Million cell updates/sec

Title: >US-08-231-565A-32  
 Description: (1-10) from US08231565A.pap  
 Perfect Score: 81  
 Sequence: 1 VLRYGFSV 10

Scoring table: PAM 150  
 Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-genseq26  
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
 14:part14 15:part15 16:part16 17:part17 18:part18  
 19:part19

Statistics: Mean 17.944; Variance 59.231; scale 0.303

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	81	100.0	10 15	R82168	Melanoma-specific mut	7.64e-02
2	81	100.0	10 15	R84200	gpi100 melanoma antige	7.64e-02
3	81	100.0	661 14	R78646	Melanoma associated a	7.64e-02
4	81	100.0	661 15	R84854	MART-1 melanoma antig	7.64e-02
5	81	100.0	661 15	R84855	MART-1 melanoma antig	7.64e-02
6	61	75.3	705 6	R31592	Prolylendopeptidase.	9.92e+00
7	61	75.3	705 15	R90712	Prolyl-endopeptidase.	9.92e+00
8	56	69.1	560 19	W05148	Human brain sodium-de	3.16e+01
9	53	65.4	555 8	R43340	Alpha-glucosidase fro	6.25e+01
10	53	65.4	561 16	R91065	Rat interleukin-1 typ	6.25e+01
11	52	64.2	409 3	R13118	Shiga-like toxin subu	7.83e+01
12	51	63.0	748 10	R52706	Human enkephalinase.	9.78e+01
13	51	63.0	748 1	R80969	Enkephalinase (human)	9.78e+01
14	51	63.0	750 1	P90393	Human common acute ly	9.78e+01
15	51	63.0	750 1	P82940	Atrial natriuretic po	9.78e+01
16	50	61.7	124 15	R75570	VH Fab 3B3 binds to g	1.22e+02
17	50	61.7	124 15	R75572	VH Fab 3B9 binds to g	1.22e+02
18	50	61.7	124 15	R75561	VH Fab M556-13 binds	1.22e+02
19	50	61.7	124 15	R75613	VH Fab M556-10 binds	1.22e+02
20	50	61.7	124 15	R75615	VH Fab M556-16 binds	1.22e+02

21	50	61.7	124 15	R75616	VH Fab M556-5 binds t	1.22e+02
22	50	61.7	124 15	R75614	VH Fab M556-15 binds	1.22e+02
23	50	61.7	124 15	R75610	VH Fab M556-2 binds t	1.22e+02
24	50	61.7	124 15	R75612	VH Fab M556-7 binds t	1.22e+02
25	50	61.7	124 15	R75611	VH Fab M556-3 binds t	1.22e+02
26	50	61.7	537 18	R95585	Atrial natriuretic pe	1.22e+02
27	50	61.7	537 2	P70432	Bovine atrial natriur	1.22e+02
28	50	61.7	541 2	P70433	Human atrial natriure	1.22e+02
29	50	61.7	541 18	R95586	Atrial natriuretic pe	1.22e+02
30	50	61.7	772 13	R70690	Mesquite allergenic p	1.22e+02
31	49	60.5	255 4	R22040	Short-form rat enkeph	1.52e+02
32	49	60.5	515 2	P70419	Chimeric cytochrome-P	1.52e+02
33	49	60.5	516 3	P61362	Soybean glycinin A3B4	1.52e+02
34	49	60.5	519 7	R34881	Chimeric cytochrome P	1.52e+02
35	49	60.5	519 2	P70576	Chimeric cytochrome P	1.52e+02
36	49	60.5	750 10	R52707	Rat enkephalinase.	1.52e+02
37	49	60.5	750 1	P82867	Enkephalinase (rat).	1.52e+02
38	49	60.5	870 13	R75410	B. sphaericus mtx tox	1.52e+02
39	49	60.5	870 3	R14529	Mosquitotoxin.	1.52e+02
40	49	60.5	870 8	R41020	Insecticidal protein	1.52e+02
41	48	59.3	357 18	R94502	Polyhydroxyacid synth	1.89e+02
42	48	59.3	382 7	R36677	Deduced from Mycobact	1.89e+02
43	48	59.3	573 19	W01869	Transglutaminase from	1.89e+02
44	48	59.3	834 1	R04869	Alpha-1-6-glucanase.	1.89e+02
45	48	59.3	910 17	R91737	HER4-Ig fusion protei	1.89e+02

## ALIGNMENTS

RESULT 1  
 ID R82168 standard; peptide; 10 AA.  
 AC R82168;  
 DT 25-MAR-1996 (first entry)  
 DE Melanoma-specific mutant immunogen epitope 10mer peptide.  
 KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
 KW Cytotoxic T cell; lymphocyte; HLA-A2.  
 OS Homo sapiens.  
 PN N09522561-A2.  
 PD 24-AUG-1995.  
 PF 16-FEB-1995; U01991.  
 PR 16-FEB-1994; US-197399.  
 PR 25-APR-1994; US-234784.  
 PA (UYVI-) UNIV VIRGINIA PATENT FOUND.  
 PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
 DR WPI: 95-302688/39.  
 PT Melanoma-specific immunogen comprises epitope(s) homologous with  
 PT pMEL17 - are highly potent stimulators of HLA-A2+CTL's useful in  
 PT adoptive immuno-therapy  
 PS Example 8; Page 52; 148pp; English.  
 CC A melanoma-specific immunogen homologous with pMEL-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The immunogen can be used for  
 CC R82108 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pMEL-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 81; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7.64e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 VLRYGFSV 10  
 |||||  
 QY 1 VLRYGFSV 10

RESULT 2  
 ID R84200 standard; Peptide; 10 AA.  
 AC R84200;  
 DT 25-APR-1996 (first entry)  
 DE gpi100 melanoma antigen immunogenic peptide (G10-5).

KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WP9529493-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 55; Page 131; 184pp; English.  
 CC The immunogenic peptide is derived from cDNA25 (R84854), a  
 CC melanoma antigen derivative of gp100 (see R84855). The  
 CC peptide and its derivatives (see R84200-R84211) are used in  
 CC medicaments (vaccines) for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against melanoma-specific  
 CC antigens and its immunogenic peptides may be used in the  
 CC detection and isolation of the antigen from a sample, the  
 CC detection of which is indicative of a disease state  
 CC (melanoma or metastatic melanoma).  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 81; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 7.64e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vlyrygsfsv 10  
 |||||  
 Qy 1 VLIRYGSFSV 10

RESULT 3  
 ID R78646 standard; Protein; 661 AA.  
 AC R78646;

DE 22-JAN-1996 (first entry)  
 DE Melanoma associated antigen gp100.  
 KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 OS Homo sapiens.

PN EP-668350-A1.  
 PD 23-AUG-1995.  
 PF 14-FEB-1997; 200348.  
 PR 16-FEB-1994; EP-200337.  
 PR 21-DEC-1994; EP-203709.  
 PA (ALKU ) ARZO NOBEL NV.  
 PI Adema GJ, Figdor CG;  
 DR WPI; 95-284790/38.  
 DR N-PSDB; 096055.  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PT detection of tumours  
 PS Claim 1; Page 22-24; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC may be used in the production of vaccines. Nucleotide sequences  
 CC encoding the immunogenic peptides may be used as primers and probes  
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
 CC capable of binding to the melanoma associated antigen can be  
 CC cultured ex vivo and returned to melanoma particles, and when  
 CC radiolabelled, they may be used to identify tumour deposits.  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 81; DB 14; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 7.64e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 476 vlyrygsfsv 485  
 |||||  
 Qy 1 VLIRYGSFSV 10

RESULT 4  
 ID R84854 standard; Protein; 661 AA.  
 AC R84854;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen cDNA25.  
 KW cDNA25; MART-1; melanoma antigen recognised by T-cell;  
 KW gp100 antigen derivative; melanoma; metastatic melanoma;  
 KW tumour-associated antigen; immunogen; diagnosis; prognosis;  
 KW prophylaxis; therapy; vaccine.  
 OS Mammalian sp.  
 FH Key Location/Qualifiers  
 FT Peptide 457..466  
 FT /label= antigenic\_peptide  
 FT /note= "see R84199"  
 PN WO9529493-A2.  
 PD 02-NOV-1995.  
 PR 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DR N-PSDB; T02716.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 5A; 184pp; English.  
 CC cDNA2 is a melanoma antigen (MART-1) which is recognized by  
 CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-  
 CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source  
 CC of immunogenic peptides (see R84199) which are optionally modified  
 CC (see R84200-R84211) to enhance their binding to a MHC molecule and  
 CC used in medicaments, especially vaccines, for the treatment or  
 CC prevention (by immunisation) of melanoma. Antibodies against cDNA2  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of the antigen from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 81; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 7.64e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 476 vlyrygsfsv 485  
 |||||  
 Qy 1 VLIRYGSFSV 10

RESULT 5  
 ID R84855 standard; Protein; 661 AA.  
 AC R84855;

DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen gp100.  
 KW gp100; MART-1; melanoma antigen recognised by T-cell;  
 KW cDNA25 antigen derivative; melanocyte; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian sp.

FH Key Location/Qualifiers  
 FT Peptide 154..163  
 FT /label= G9-154\_immunogenic\_peptide  
 FT Peptide 208..217  
 FT /label= G9-209\_immunogenic\_peptide  
 FT /note= "see R84210"  
 FT Peptide 280..288  
 FT /label= G9-280\_immunogenic\_peptide  
 FT /note= "see R84208"  
 FT Peptide 457..266  
 FT /label= immunogenic\_peptide



FT Peptide 476..485  
 FT Aba1= Immunogenic\_peptide  
 PN MO9529193-A2.  
 PD 02-NOV-1995.  
 PF 11-APR-1995; U05063.  
 PR 23-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI: 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 7A; 184pp; English.

CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which  
 CC is recognized by T-lymphocytes, and is a derivative of the  
 CC melanoma-specific antigen CDNA25 (see R84854). gp100 is a source  
 CC of immunogenic peptides which are optionally modified to enhance  
 CC their binding to a MHC molecule, and used in medicaments,  
 CC especially vaccines, for the treatment or prevention (by  
 CC immunisation) of melanoma. Antibodies against CDNA2 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 81; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 7,64e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 476 vlyrygsfsv 485  
 |||||  
 QY 1 VLYRYGSFVS 10

RESULT 6  
 ID R31592 standard; Protein; 705 AA.

AC R31592;  
 DT 24-MAY-1993 (first entry)  
 DE Prollyl-endopeptidase.  
 KW Peptidase: proteinase.  
 OS Flavobacterium meningosepticum.  
 PN EP-524906-A.  
 PD 27-JAN-1993.  
 PF 15-JUL-1992; 810537.  
 PR 24-JUL-1991; EP-810595.  
 PR 12-MAR-1992; GB-005457.  
 PA (CIBA ) CIBA GEIGY JAPAN LTD.  
 PI Inaoka T, Kokubo T, Tsuru D, Yoshimoto T;  
 DR WPI: 93-028949/04.  
 DR N-PSDB: Q36021.

PT DNA encoding recombinant prolyl endo-peptidase - used in prodn.  
 PT of large quantities of enzyme for conversion of precursor  
 PT peptide(s) to amidated forms  
 PS Claim 15; Page 19; 34pp; English.  
 CC This sequence represents prolyl-endopeptidase from Flavobacterium  
 CC meningosepticum, and was decoded from the appropriate genomic DNA.  
 CC Recombinant prolyl-endopeptidase is useful for prodn. of  
 CC biologically active peptides whose C termini are alpha-amidated and  
 CC have proline residues at or near their C-termini. Examples include  
 CC LHRH, oxytocin, vasopressin, calcitonin, aspartocin, alpha-MSH,  
 CC and morphine modulating neuropeptide.  
 SQ Sequence 705 AA;

Query Match 75.3%; Score 61; DB 6; Length 705;  
 Best Local Similarity 50.0%; Pred. No. 9.92e+00;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 472 ilysyvggfnl 481  
 :|||  
 QY 1 VLYRYGSFVS 10

RESULT 7  
 ID R90712 standard; Protein; 705 AA.  
 AC R90712;  
 DT 06-APR-1996 (first entry)  
 DE prolyl-endopeptidase.  
 KW prolyl-endopeptidase; Flavobacterium meningosepticum; probe;  
 KW plasmid pFPEP02; Escherichia coli; thermostable; mutant;  
 KW random mutagenesis; polymerase chain reaction; PCR; PEP-227;  
 KW PEP-361; PEP-407; PEP-15; enzyme engineering; protein engineering;  
 KW plasmid pUK-FPEP-227; plasmid pUK-FPEP-361; plasmid pUK-FPEP-407;  
 KW plasmid pUK-FPEP-15; cloning; baculovirus; molecular evolution;  
 KW peptide hydrolysis; peptide coupling; biologically active peptide;  
 KW transpeptidation; C-terminal amidation; condensation.  
 KW Flavobacterium meningosepticum strain IFO 12535 (ATCC 13253).

OS Key Location/Qualifiers  
 FT Peptide 1..19  
 FT /note= "Signal peptide"  
 FT Region 28..34  
 FT /note= "Region used to construct probe A-12"  
 FT Misc\_difference 48  
 FT /note= "May be replaced by Gln (claim 10)"  
 FT Misc\_difference 51  
 FT /note= "May be replaced by Leu (claim 10)"  
 FT Misc\_difference 129  
 FT /note= "May be replaced by Thr (claim 10)"  
 FT Region 182..190  
 FT /note= "Region used to construct probe A-13"  
 FT Region 268..276  
 FT /note= "Region used to construct probe A-19"  
 FT Region 352..364  
 FT /note= "Region used to construct probe A-9"  
 FT Region 380..391  
 FT /note= "Region used to construct probe A-18"  
 FT Misc\_difference 477  
 FT /note= "May be replaced by Lys (claim 10)"  
 FT Region 499..509  
 FT /note= "Region used to construct probe A-3"  
 FT Misc\_difference 633  
 FT /note= "May be replaced by Val (claim 10)"  
 PN WO9600293-A1.

PD 04-JAN-1996.  
 PF 16-JUN-1995; IB0489.  
 PR 23-JUN-1994; EP-810377.  
 PA (CIBA ) CIBA GEIGY JAPAN LTD.  
 PA (JAPA-) JAPAT LTD.  
 PI Inaoka T, Kokubo T, Ohkuma-Soyejima T;  
 DR WPI: 96-068877/07.  
 DR N-PSDB: T15588.  
 PT New mutant heat-stable prolyl-endo-peptidase enzymes - from  
 PT F. meningosepticum. pred. by mutagenesis of encoding DNA, useful for  
 PT C-terminal amidation of biologically active peptide(s) e.g. insulin  
 PT or calcitonin  
 PS Claim 6; Page 43-49; 66pp; English.

CC The sequence represents a prolyl-endopeptidase from Flavobacterium  
 CC meningosepticum and is encoded by a gene isolated (using probes  
 CC A-3, A-9, A-12, A-13, A-18 and A-19, based on the native enzyme  
 CC sequence) in plasmid pFPEP02 (FERM BP-3466). The enzyme may be  
 CC modified for improved thermostability by molecular evolution  
 CC (enzyme engineering by random chemical mutagenesis and modified  
 CC polymerase chain reaction). Mutants E48Q (PEP-227, from plasmid  
 CC pUK-FPEP-227), E48Q/AL29T/G633V (PEP-361, from plasmid  
 CC pUK-FPEP-361), E48Q/AL29T/G633V/F51L (PEP-407, from plasmid  
 CC pUK-FPEP-407) and E477K (PEP-15, from plasmid pUK-FPEP-15) are  
 CC specified, and may be expressed in E. coli or a baculovirus system.  
 CC The enzyme catalyzes selective hydrolytic cleavage of peptides at  
 CC the C-terminal side of a proline residue, or coupling of peptides  
 CC by condensation or transpeptidation, and may also be used in  
 CC C-terminal amidation of biologically active peptides. The mutants  
 CC show improved efficiency, yield, conversion and resistance to  
 CC denaturation.  
 SQ Sequence 705 AA;

Query Match 75.3%; Score 61; DB 15; Length 705;

Best Local Similarity 50.0%; Pred. No. 9.92e+00; Mismatches 4; Indels 0; Gaps 0;

Matches 5; Conservative 4;

Db 472 ilysyggfni 481  
 QY 1 VLYRGSFSV 10

RESULT 8

ID W05148 standard; Protein; 560 AA.  
 AC W05148;  
 DT 30-JAN-1997 (first entry)  
 DE Human brain sodium-dependent inorganic phosphate cotransporter.  
 KW Brain sodium-dependent inorganic phosphate cotransporter; hBNPI.  
 OS Homo sapiens.  
 PN W09634288-A1.  
 PD 31-OCT-1996.  
 PF 25-APR-1996; U05792.  
 PR 27-APR-1995; US-430033.  
 PA (ELIL ) LILLY & CO ELI.  
 PI Ni B, Paul SM;  
 DR WPI: 96-497773/49.  
 DR N-PSDB; T42064.  
 PT New isolated human brain sodium-dependent inorganic phosphate  
 PT co-transporter - used to develop prods for diagnosis, treatment and  
 PT prevention of conditions involving inappropriate stimulation  
 PS Claim 1: Page 55-57; 68pp; English.  
 CC A novel human brain sodium-dependent inorganic phosphate  
 CC cotransporter (W05148), designated hBNPI, is selectively expressed  
 CC in discrete populations of neurons and glia. Its amino acid  
 CC sequence was deduced from a cDNA clone (T42064) isolated from a  
 CC human hippocampus cDNA library. hBNPI can be expressed in  
 CC transformed host cells and used to determine the effectiveness  
 CC of test cpds. for the treatment or prevention of disorders  
 CC associated with an inappropriate stimulation of hBNPI.  
 SQ Sequence 560 AA;

Query Match 69.1%; Score 56; DB 19; Length 560;  
 Best Local Similarity 60.0%; Pred. No. 3.16e+01;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 237 vfyvysgfi 246  
 QY 1 VLYRGSFSV 10

RESULT 9

ID R43340 standard; Protein; 555 AA.  
 AC R43340;  
 DT 05-APR-1994 (first entry)  
 DE Alpha-glucosidase from Bacillus stearothermophilus.  
 KW Alpha-amylase; contamination; starch; diagnosis; pancreatic;  
 KW salivary gland; diseases.  
 OS Bacillus stearothermophilus ATCC 12016.  
 FH Key Location/Qualifiers  
 FT Peptide 1..15  
 FT /note= "N-terminal fragment"  
 PN EP-567785-A.  
 PD 03-NOV-1993.  
 PF 26-MAR-1993; 105023.  
 PR 27-MAR-1992; JP-101658.  
 PR 11-MAY-1992; JP-117538.  
 PA (TOYM ) TOYO BOSEKI KK.  
 PI Emi S, Nishiya Y, Sogabe A, Sogabe Y, Suzuki Y;  
 PI Takii Y, Yamamoto K;  
 DR WPI: 93-346249/44.  
 DR N-PSDB; Q51151.  
 PT Alpha-glucosidase protein and DNA encoding it - obtd. from  
 PT Bacillus stearothermophilus ATCC 12016, used for starch prodn.  
 PS Claim 1: Page 9-12; 16pp; English.  
 CC The alpha-glucosidase gene was isolated from the chromosomal DNA of  
 CC Bacillus stearothermophilus ATCC 12016. The gene allows production

CC of pure alpha glucosidase by recombinant techniques, free of alpha-  
 CC amylase contamination and with a greater specific activity than  
 CC previously found. Alpha-glucosidase is used for the production of  
 CC starches or for the determination of the amt. of alpha-amylase  
 CC present in body fluids for the diagnosis of pancreatic or salivary  
 CC gland diseases.  
 SQ Sequence 555 AA;

Query Match 65.4%; Score 53; DB 8; Length 555;  
 Best Local Similarity 71.4%; Pred. No. 6.25e+01;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 512 lyrydgf 518  
 QY 2 LYRGSF 8

RESULT 10

ID R91065 standard; Protein; 561 AA.  
 AC R91065;  
 DT 23-MAY-1996 (first entry)  
 DE Rat interleukin-1 type-3 receptor.  
 KW Interleukin-1 type-3 receptor; IL-1-3R; immune-associated disease;  
 KW vector; antibody; therapy.  
 OS Rattus sp.  
 FH Key Location/Qualifiers  
 FT Domain 1..338  
 FT /label= Extracellular\_domain  
 FT Domain 339..359  
 FT /label= Transmembrane\_domain  
 FT Domain 360..561  
 FT /label= Intracellular\_domain  
 PN W09607739-A2.  
 PD 14-MAR-1996.  
 PF 11-SEP-1995; U12037.  
 PR 09-SEP-1994; US-303957.  
 PA (NEUR-) NEUROCRINE BIOSCIENCES INC.  
 PI Clevenger W, DeSouza EB, Liaw CW, Lovenberg TW;  
 PI Oltersdorf T;  
 DR WPI: 96-171614/17.  
 DR N-PSDB; T13167.

PT Interleukin-1 type 3 receptor proteins - useful for the treatment of  
 PT immune-associated diseases  
 PS Claim 5: Page 49-52; 64pp; English.  
 CC Rat interleukin-1 type-3 receptor (IL-1-3R) (R91065) is a protein  
 CC the binds IL-1 (alpha or beta) and, when expressed on a cell surface,  
 CC transduces the signal provided by IL-1 to the cell. It can be  
 CC obtd. by expression in prokaryotic or eukaryotic host cells of an  
 CC encoding cDNA sequence (T13167) isolated from a rat lung cDNA library.  
 CC IL-1-3R or its soluble form (extracellular domain only) is used as  
 CC a therapeutic reagent, immunogen, immunoassay reagent or affinity  
 CC ligand, or is used to screen IL-1-3R (ant)agonists, or to raise  
 CC antibodies useful for flow cytometry.  
 SQ Sequence 561 AA;

Query Match 65.4%; Score 53; DB 16; Length 561;  
 Best Local Similarity 50.0%; Pred. No. 6.25e+01;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 355 ilviyntfky 364  
 QY 1 VLYRGSFSV 10

RESULT 11

ID R13118 standard; Protein; 409 AA.  
 AC R13118;  
 DT 08-OCT-1991 (first entry)  
 DE Shiga-like toxin subunit A.  
 KW Hybrid; fusion; membrane translocation; binding region; HIV;  
 KW infection; toxin; steroid; hormone; monoclonal antibody; antigen;  
 KW diphtheria; exotoxin; phenylketonuria; cholera; interleukin; IL-2;  
 KW protease; epidermal growth factor; ricin; tetanus; hexosaminidase;

KW Shiga-like toxin A; SLT-A; ligand; insulin; nuclease.  
 OS Vibrio cholera.  
 FH Key  
 FT Misc\_difference 317  
 FT /note= "encoded by stop codon"  
 PN WO9109871-A.  
 PD 11-JUL-1991.  
 PF 21-DEC-1990; U07619.  
 PR 22-DEC-1989; US-456095.  
 PR 14-JUN-1990; US-538276.  
 PA (SERA.) SERAGEN INC.  
 PI Murphy JR.  
 DR WPI; 91-222845/30.  
 DR N-PSDB; Q12711.  
 FT Hybrid molecules for targetting chemical entity to cell - have  
 FT membrane trans-locating and cell binding-regions and used to  
 FT treat HIV infection, genetic enzyme-deficiency disorders etc.  
 PS Disclosure; Fig 8(1-3); 59pp; English.  
 CC Hybrid molecules are produced by covalently linking  
 CC (1) a portion (A) of the binding domain of a cell-binding ligand,  
 CC allowing binding of the mol. to an animal cell;  
 CC (2) a portion (B) of a translocation domain of a protein able to  
 CC translocate (C) across the cell cytoplasmic membrane, and  
 CC and (3) a portion (C) which is to be introduced into the cell.  
 CC (A) is derived from a steroid or polypeptide hormone, a single-chain  
 CC analogue of a monoclonal antibody able to bind an antigen expressed  
 CC on the cell surface, or a polypeptide toxin.  
 CC (B) is derived from a toxin (e.g. diphtheria toxin or Pseudomonas  
 CC exotoxin A).  
 CC (A) may be derived from insulin, interleukins 2, 3 or 6 or  
 CC epidermal growth factor.  
 CC Suitable enzymes in (C) include cholera toxin, ricin, tetanus toxin,  
 CC hexosaminidase A, protease, nuclease, SLT-A, etc.  
 CC Specified examples are CT-A/DT-B'/IL-2, SLTA/DT-B'/IL-2,  
 CC ricin A/DT-B'/IL-2, HIVP-BP/DT-B'/IL-2 and the phenylalanine  
 CC hydroxylase-PT-B' or their biologically active mutants.  
 CC (CT-A= cholera toxin, DT-B'= truncated diphtheria toxin,  
 CC SLTA= Shiga-like toxin A; HIVP-BP= HIV protease binding protein.  
 CC See also Q12710-12.  
 CC Sequence 409 AA;  
 SQ

Query Match 64.2%; Score 52; DB 3; Length 409;  
 Best Local Similarity 55.6%; Pred. No. 7.83e+01;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 110 vfyrfadfs 118  
 QY 1 VLYRYGFS 9  
 :|:|:|:|

RESULT 12  
 ID R52706 standard; Protein; 748 AA.  
 AC R52706;  
 DT 05-JAN-1995 (first entry)  
 DE Human enkephalinase.  
 KW Human; enkephalinase.  
 KW kidney brush border neutral proteinase; pituitary; brain; lymph nodes;  
 KW neurophils; enkephalins; tachykinins.  
 OS Homo sapiens.  
 FH Key  
 FT Protein  
 FT /note= "Mature enkephalinase"  
 FT Peptide  
 FT /note= "Stop transfer sequence - the conformational  
 FT bends introduced by the Pro residues and the  
 FT electropositive character provided by the Lys  
 FT residues act together with the transmembrane  
 FT region to bar transfer of enkephalinase through  
 FT the cell membrane"  
 PN EP-596355-A.  
 PD 11-MAY-1994.  
 PF 23-DEC-1987; 117230.  
 PR 24-DEC-1986; US-946566.

Query Match 64.2%; Score 52; DB 3; Length 409;  
 Best Local Similarity 55.6%; Pred. No. 7.83e+01;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 110 vfyrfadfs 118  
 QY 1 VLYRYGFS 9  
 :|:|:|:|

Query Match 63.0%; Score 51; DB 10; Length 748;  
 Best Local Similarity 55.6%; Pred. No. 9.78e+01;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 620 mvyqgnfs 628  
 QY 1 VLYRYGFS 9  
 :|:|:|:|

RESULT 13  
 ID P80969 standard; protein; 748 AA.  
 AC P80969;  
 DT 15-NOV-1990 (first entry)  
 DE Enkephalinase (human).  
 KW Enkephalinase; immunosuppressant.  
 OS Homo sapiens.  
 PN EP-272928-A.  
 PD 29-JUN-1988.  
 PF 23-DEC-1987.  
 PR 24-DEC-1986; US-946566.  
 PR 12-JAN-1987; US-002478.  
 PA (GETH) Genentech Inc.  
 PI Malfroy-Camine B, Scofield PR;  
 DR WPI; 88-177004/26.  
 PT Deoxyribonucleic acid isolates coding for enkephalinase - and prodn.  
 PT of enkephalinase, used in therapeutic applications in diagnostic  
 PT immunoassay and as immunosuppressants.  
 PS Disclosure; p; English.  
 CC The DNA was sequenced from clones isolated from a lambda gt10  
 CC library constructed from poly(A+) human placental RNA. At posn.  
 CC 1413 one clone, lambdaH7 was observed to have a G (coding for Ala);  
 CC in clone lambdaH5 this was found to be an A (Thr). Since the former  
 CC is identical to the rat amino acid at posn. 465, the latter  
 CC probably represents an error of the reverse transcriptase synthesis  
 CC of the mRNA. The DNA can be ligated into an expression vector for  
 CC prodn. of the enkephalinase protein in a host. The recombinant  
 CC orders associated with endogenous peptides such as tachykinins and  
 CC kinins, eg. acute inflammation, and hyperimmune responses. It is  
 CC also useful in diagnostic immunoassays, and as immunosuppressants  
 CC by virtue of their ability to digest chemotactic molecules.  
 CC This sequence is also published in EP-272929.  
 CC See also P82867.  
 CC Sequence 748 AA;  
 SQ

Query Match 63.0%; Score 51; DB 1; Length 748;  
 Best Local Similarity 55.6%; Pred. No. 9.78e+01;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 620 mvyqgnfs 628  
 QY 1 VLYRYGFS 9  
 :|:|:|:|

PR 12-JAN-1987; US-002478.  
 PA (GETH ) GENENTECH INC.  
 PI Malfroy-Camine B, Schofield PR;  
 DR WPI; 94-152785/19.  
 DR N-PSDB; Q62678.  
 PT Method for assay of enkephalinase using dansylated peptide(s) -  
 PT for identification of recombinant enkephalinase  
 PS Disclosure; Fig 1; 61pp; English.  
 CC This sequence represents human enkephalinase. Enkephalinase is a  
 CC neutral endopeptidase or kidney brush border neutral proteinase. It  
 CC has been isolated from the kidney, intestine, pituitary, brain and in  
 CC lymph nodes, and has been detected in many peripheral organs and in  
 CC human neutrophils. The distribution of enkephalinase in the brain  
 CC closely parallels the distribution of the enkephalins. Mammalian  
 CC enkephalinase may be used in the treatment of various pathological  
 CC disorders associated with various endogenous peptides, eg.  
 CC tachykinins and kinins.  
 SQ Sequence 748 AA;

Query Match 63.0%; Score 51; DB 10; Length 748;  
 Best Local Similarity 55.6%; Pred. No. 9.78e+01;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 Db 620 mvyqgnfs 628  
 QY 1 VLYRYGFS 9  
 :|:|:|:|

Query Match 63.0%; Score 51; DB 1; Length 748;  
 Best Local Similarity 55.6%; Pred. No. 9.78e+01;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 620 mvyqgnfs 628  
 QY 1 VLYRYGFS 9  
 :|:|:|:|

## RESULT 14

ID P90393 standard; protein; 750 AA.  
 AC P90393;  
 DT 1-NOV-1989 (first entry)  
 DE Human common acute lymphoblastic leukaemia antigen  
 KW Deduced human common acute lymphoblastic leukaemia antigen;  
 KW analgesic; inhibits leukaemia; endopeptidase.  
 OS Homo sapiens (Human)  
 PN WO8905353-A.  
 PD 15-JUN-1989.  
 PF 01-DEC-1988; U04280.  
 PR 04-DEC-1987; US-247915.  
 PA (DANA) Dana Farber Cancer Inst.  
 PI Reinherz EL, Shipp MA, Richardson NE, Ritz J, Sayre PH;  
 DR WPI; 89-192699/26.  
 DR N-PSDB; N90123.  
 PT DNA encoding human common acute lymphoblastic leukaemia antigen  
 PT - used for obtaining pure protein for diagnosis and  
 PT treatment of medical conditions.  
 PS Claim 4; page 49 and fig 3; 69pp; English.  
 CC Deduced amino acid sequence of human common acute lymphoblastic  
 CC leukaemia antigen (CALLA) and its fragments. Binds to  
 CC anti-CALLA antibody, and is used for diagnosis, treatment, to  
 CC design a drug for use as a general analgesic. Has functional  
 CC endopeptidase activity.  
 SQ Sequence 750 AA;

Query Match 63.0%; Score 51; DB 1; Length 750;

Best Local Similarity 55.6%; Pred. No. 9.78e+01;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 622 mvyyqygnfs 630

:::|::|

QY 1 VLRYGSFS 9

## RESULT 15

ID P82940 standard; protein; 750 AA.  
 AC P82940;  
 DT 23-NOV-1990 (first entry)  
 DE Atrial natriuretic polypeptide binding polypeptide.  
 KW Diuretic; atrium cardis; atrial natriuretic peptide binding;  
 KW hypotensive action.  
 PN J63079598-A.  
 PD 09-APR-1988.  
 PF 22-SEP-1986; 222192.  
 PR 22-SEP-1986; JP-222192.  
 PA (SUNR) Suntory Ltd.  
 DR WPI; 88-137132/20.  
 DR N-PSDB; N81690.  
 PT Novel polypeptide, with diuretic action - is obtd. from atrium cardis  
 PT of mammals showing specific bond to atrial natriuretic polypeptide  
 PT and gene codinh it.  
 PS Disclosure; p; Japanese.  
 CC Peptide specifically binds to atrial natriuretic polypeptide (ANP)  
 CC and has diuretic (partic. natriuretic) and hypotensive action.  
 CC See also N81690-93 and P81282-P81309.  
 SQ Sequence 750 AA;

Query Match 63.0%; Score 51; DB 1; Length 750;

Best Local Similarity 55.6%; Pred. No. 9.78e+01;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 622 mvyyqygnfs 630

:::|::|

QY 1 VLRYGSFS 9

Search completed: Tue Jun 10 11:23:37 1997

Job time : 9 secs.



```
ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS
REFERENCE     A41234
#authors      Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
              Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
              Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
              coat color locus on mouse chromosome 10 and is in a
              syntenic region on human chromosome 12.
#cross-references MUID:92021023
#accession    A41234
#status       preliminary
#molecule_type mRNA
#residues     1-668 #label KWO
#cross-references GB:M77348
SUMMARY       #length 668 #molecular-weight 70932 #checksum 6409
Query Match   100.0%; Score 81; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 4.58e-04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 476 vlyrygsfsv 485
QY 1 VLYRYGSFSV 10
RESULT 3
ENTRY      A49179 #type fragment
TITLE      melanoma antigen homolog rpe1 - bovine (fragment)
ORGANISM   #formal_name Bos primigenius taurus #common_name cattle
DATE       19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
ACCESSIONS A49179
REFERENCE   A49179
#authors    Kim, R.Y.; Wistow, G.J.
#journal    Exp. Eye Res. (1992) 55:657-662
#title      The cDNA RPE1 and monoclonal antibody HMB-50 define gene
              products preferentially expressed in retinal pigment
              epithelium.
#cross-references MUID:93122163
#accession  A49179
#status     preliminary
#molecule_type nucleic acid
#residues   1-491 #label KIM
#cross-references NCBI:122438; NCBI:P:122439
#experimental_source retinal pigment epithelium
#note       #length 491 #checksum 3125
SUMMARY     #length 491 #checksum 3125
Query Match   92.6%; Score 75; DB 14; Length 491;
Best Local Similarity 90.0%; Pred. No. 5.92e-03;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 305 vlyrygsfsv 314
QY 1 VLYRYGSFSV 10
RESULT 4
ENTRY      S53871 #type complete
TITLE      Pmel 17 protein - mouse
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE       27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
ACCESSIONS S53871
REFERENCE   S53871
#authors    Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
              Chintamaneni, C.; Bennett, D.; Pickard, R.T.
#journal    Nucleic Acids Res. (1995) 23:154-158
#title      Mouse silver mutation is caused by a single base insertion in
```

```
#accession    S53871
#status       preliminary
#molecule_type mRNA
#residues     1-626 #label KWO
SUMMARY       #length 626 #molecular-weight 65979 #checksum 5710
Query Match   92.6%; Score 75; DB 14; Length 626;
Best Local Similarity 90.0%; Pred. No. 5.92e-03;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 443 vlyrygsfsv 452
QY 1 VLYRYGSFSV 10
RESULT 5
ENTRY      JX0194 #type complete
TITLE      prolyl oligopeptidase (EC 3.4.21.26) - Flavobacterium
              meningosepticum
ALTERNATE_NAMES prolyl endopeptidase
ORGANISM         #formal_name Flavobacterium meningosepticum
DATE             31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
ACCESSIONS       JX0194; PS0235; S19201
REFERENCE         JX0194
#authors          Yoshimoto, T.; Kanatani, A.; Shimoda, T.; Inaoka, T.; Kokubo,
              T.; Tsuru, D.
#journal          J. Biochem. (1991) 110:873-878
#title            Prolyl endopeptidase from Flavobacterium meningosepticum:
              cloning and sequencing of the enzyme gene.
#cross-references MUID:92176159
#accession        JX0194
#molecule_type  DNA
#residues         1-705 #label YOS2
#accession        PS0235
#molecule_type  protein
#residues         33-53; 73-78; 82-85; 113-122; 155-162; 190-197; 201-209;
              223-229; 286-302; 307-313; 316-319; 365-383; 399-410;
              433-441; 518-525; 551-557; 584-589; 600-605; 609-614;
              690-698 #label YOS
REFERENCE         S19197
#authors          Diefenthal, T.
#submission       Submitted to the EMBL Data Library, January 1992
#accession        S19201
#status           preliminary
#molecule_type  DNA
#residues         1-109; 'C', 111-586, 'A', 588-705 #label DIE
#cross-references EMBL:X63674
CLASSIFICATION   #superfamily prolyl oligopeptidase
KEYWORDS          hydrolase; serine proteinase
FEATURE           556
SUMMARY           #active_site Ser #status experimental
              #length 705 #molecular-weight 78706 #checksum 7699
Query Match     75.3%; Score 61; DB 5; Length 705;
Best Local Similarity 50.0%; Pred. No. 1.63e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 472 ilysyggfsv 481
QY 1 VLYRYGSFSV 10
RESULT 6
ENTRY      A38086 #type complete
TITLE      prolyl oligopeptidase (EC 3.4.21.26) precursor -
              Flavobacterium meningosepticum
ORGANISM         #formal_name Flavobacterium meningosepticum
DATE             07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change
ACCESSIONS       A38086
REFERENCE         A38086
```

```

#authors      Chevallier, S.; Goeltz, P.; Thibault, P.; Banville, D.;
#journal      J. Biol. Chem. (1992) 267:8192-8199
#title       Characterization of a prolyl oligopeptidase from
#           Flavobacterium meningosepticum. Complete sequence and
#           localization of the active-site serine.
#accession   A38086
##status     preliminary
##molecule_type DNA
##residues   1-705 ##label CHE
##cross-references GB:M81461
CLASSIFICATION #superfamily prolyl oligopeptidase
KEYWORDS       hydrolase; serine proteinase
SUMMARY        #length 705 #molecular-weight 78839 #checksum 6349

Query Match      75.3%; Score 61; DB 5; Length 705;
Best Local Similarity 50.0%; Pred. No. 1.63e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 472 llysgggnf 481
:|:|:|:|:|
Qy 1 VLRYGFSV 10

RESULT 7
ENTRY   JN0585 #type complete
TITLE   prolyl oligopeptidase (EC 3.4.21.26) - Aeromonas hydrophila
        (strain JM83)
ALTERNATE_NAMES prolyl endopeptidase
ORGANISM #formal_name Aeromonas hydrophila
DATE 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
JN0585; PN0498
ACCESSIONS JN0585; PN0498
REFERENCE JN0585
#authors Kanatani, A.; Yoshimoto, T.; Kitazono, A.; Kokubo, T.; Tsuru,
D
#journal J. Biochem. (1993) 113:790-796
#title Prolyl endopeptidase from Aeromonas hydrophila: Cloning,
        sequencing, and expression of the enzyme gene, and
        characterization of the expressed enzyme.
#accession JN0585
##molecule_type DNA
##residues 1-689 ##label KAN
##cross-references DDBJ:D14005
#accession PN0498
##molecule_type protein
##residues 18-51;62-71;149-157;190-215;218-223;297-316;461-467;
        511-526;565-578;581-587;593-608;639-653 ##label KA2
CLASSIFICATION #superfamily prolyl oligopeptidase
KEYWORDS hydrolase; serine proteinase
FEATURE 2-689 #product prolyl oligopeptidase #label POP\
512,537,656 #active_site Asp, Ser, His #status predicted
SUMMARY #length 689 #molecular-weight 76384 #checksum 7164

Query Match      71.6%; Score 58; DB 5; Length 689;
Best Local Similarity 60.0%; Pred. No. 5.02e+00;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 453 llygggfdv 462
:|:|:|:|
Qy 1 VLRYGFSV 10

RESULT 8
ENTRY   S51251 #type complete
TITLE   probable membrane protein YDR100w - yeast (Saccharomyces
        cerevisiae)
ALTERNATE_NAMES hypothetical protein YD8557.09
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 23-Feb-1995 #sequence_revision 12-May-1995 #text_change
23-Aug-1996
ACCESSIONS S51251

```

```

REFERENCE S51243
#authors Murphy, L.; Harris, D.
#submission submitted to the EMBL Data Library, January 1995
#accession S51251
##molecule_type DNA
##residues 1-143 ##label MUR
##cross-references EMBL:Z47746
GENETICS
#map_position 4R
KEYWORDS transmembrane protein
FEATURE 9-25 #domain transmembrane #status predicted #label TM1\
37-53 #domain transmembrane #status predicted #label TM2\
76-92 #domain transmembrane #status predicted #label TM3\
94-110 #domain transmembrane #status predicted #label TM4
SUMMARY #length 143 #molecular-weight 15881 #checksum 2983

Query Match      70.4%; Score 57; DB 11; Length 143;
Best Local Similarity 85.7%; Pred. No. 7.24e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 62 lryasf 68
:|:|:|:|
Qy 2 LYRYGSF 8

RESULT 9
ENTRY   S53933 #type complete
TITLE   probable membrane protein YGR210c - yeast (Saccharomyces
        cerevisiae)
ALTERNATE_NAMES hypothetical protein G7748
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change
19-Jul-1996
ACCESSIONS S53933; S61948; S64532
REFERENCE S53922
#authors Guerreiro, P.; Barreiros, T.; Soares, H.; Cyrne, L.; Maia e
        Silva, A.; Rodrigues-Pousada, C.
#submission submitted to the EMBL Data Library, April 1995
#description Sequencing of a 17.6 kb segment on the right arm of yeast
        Chromosome VII reveals 12 open reading frames, including
        CCT, ADE3 and TR-I genes, homologous to the yeast YAL023
        and EFTG genes, of the human.
#accession S53933
##molecule_type DNA
##residues 1-411 ##label GUE
##cross-references EMBL:Z49133
##experimental_source strain S288C
REFERENCE S61947
#authors Song, J.M.; Cheung, E.; Rabinowitz, J.C.
#submission submitted to the EMBL Data Library, November 1995
#description Analysis of the 15.6-kb fragment encompassing the ADE3 gene.
#accession S61948
##molecule_type DNA
##residues 1-411 ##label SON
##cross-references EMBL:U40843
##experimental_source strain GRF88
REFERENCE S64517
#authors Guerreiro, P.; Barreiros, T.; Cyrne, L.; Soares, H.; Maia e
        Silva, A.; Rodrigues-Pousada, C.
#submission submitted to the Protein Sequence Database, May 1996
#accession S64532
##molecule_type DNA
##residues 1-411 ##label GUV
##cross-references EMBL:Z72995
##experimental_source strain S288C
GENETICS
#map_position 7R
KEYWORDS transmembrane protein
FEATURE 317-333
SUMMARY #length 411 #molecular-weight 45214 #checksum 8815

```

```

Query Match      70.4%; Score 57; DB 12; Length 411;
Best Local Similarity 70.0%; Pred. No. 7.24e+00;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 308 vlyrfgstgv 317
||||:|:|:|
QY 1 VLYRYGFSV 10

RESULT 10
ENTRY E37202 #type complete
TITLE xylan 1,4-beta-xylosidase (EC 3.2.1.37) (XynB) - Caldocellum
saccharolyticum
ORGANISM #formal_name Caldocellum saccharolyticum
DATE 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change
17-Feb-1994
ACCESSIONS E37202
REFERENCE A37202
#authors Luethi, E.; Love, D.R.; McAnulty, J.; Wallace, C.; Caughey,
P.A.; Saul, D.; Bergquist, P.L.
#journal Appl. Environ. Microbiol. (1990) 56:1017-1024
#title Cloning, sequence analysis, and expression of genes encoding
xylan-degrading enzymes from the thermophile "Caldocellum
saccharolyticum".
#cross-references MUID:90253140
#accession E37202
#status Preliminary
#molecule_type DNA
#residues 1-488 #label LUE
#cross-references GB:M34459
KEYWORDS glycosidase; hydrolase
SUMMARY #length 488 #molecular-weight 56365 #checksum 328

Query Match      70.4%; Score 57; DB 8; Length 488;
Best Local Similarity 44.4%; Pred. No. 7.24e+00;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 181 lykyaafai 189
||||:|:|:|
QY 2 LYRYGFSV 10

RESULT 11
ENTRY I59302 #type complete
TITLE brain specific Na+-dependent inorganic phosphate
cotransporter - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
02-Jul-1996
ACCESSIONS I59302
REFERENCE I59302
#authors Ni, B.; Rosteck, P.R.; Nadi, N.S.; Paul, S.M.
#journal Proc. Natl. Acad. Sci. U.S.A. (1994) 91:5607-5611
#title Cloning and expression of a cDNA encoding a brain-specific Na
(+)-dependent inorganic phosphate cotransporter.
#cross-references MUID:94261635
#accession I59302
#status Preliminary; translated from GB/EMBL/DDBJ
#molecule_type mRNA
#residues 1-560 #label RES
#cross-references EMBL:U07609; NID:g507414; CDS_PTD:g507415
SUMMARY #length 560 #molecular-weight 61665 #checksum 1275

Query Match      69.1%; Score 56; DB 14; Length 560;
Best Local Similarity 60.0%; Pred. No. 1.04e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 237 vfyvysfgl 246
|:|:|:|:|:|
QY 1 VLYRYGFSV 10

RESULT 12

```

```

ENTRY JX0190 #type complete
TITLE Cytochrome P450 1A2 - golden hamster
ALTERNATE_NAMES cytochrome P450 MC4; cytochrome P450-H (2,3,4,7,
8-pentachlorodibenzofuran inducible)
ORGANISM #formal_name Mesocricetus auratus #common_name golden hamster
DATE 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change
17-May-1996
ACCESSIONS JX0190; S13885; PX0036
REFERENCE JX0189
#authors Sagami, I.; Ohmachi, T.; Fujii, H.; Kikuchi, H.; Watanabe, M.
#journal J. Biochem. (1991) 110:641-647
#title Hamster cytochrome P-450 IA gene family, P-450 IAL and P-450
IA2 in lung and liver: cDNA cloning and sequence analysis.
#cross-references MUID:92138673
#accession JX0190
#molecule_type mRNA
#residues 1-513 #label SAG
#experimental_source lung and liver, microsomes
REFERENCE S13884
#authors Lai, T.S.; Chiang, J.Y.L.
#journal Arch. Biochem. Biophys. (1990) 283:429-439
#title Cloning and characterization of two major
3-methylcholanthrene inducible hamster liver cytochrome
P450s.
#cross-references MUID:91112759
#accession S13885
#molecule_type mRNA
#residues 1-48, 'F', 50-51, 'MC', 54-252, 'GG', 255-325, 'W', 327-355, 'L',
357-484, 'O', 486-513 #label LAI
#cross-references EMBL:M63787
REFERENCE PX0036
#authors Koga, N.; Ariyoshi, N.; Nakashima, H.; Yoshimura, H.
#journal J. Biochem. (1990) 107:826-833
#title Purification and characterization of two forms of 2,3,4,7,
8-pentachlorodibenzofuran-inducible cytochrome P-450 in
hamster liver.
#cross-references MUID:90361684
#accession PX0036
#molecule_type protein
#residues 2-19 #label KOG
#experimental_source liver
COMMENT Cytochrome P-450 I family consists of two members, IAL and IA2.
Both of them are inducible by 3-methylcholanthrene and 2,3,7,
8-tetrachlorodibenzo-p-dioxin, but have different substrate
specificities.
GENETICS
#gene CYP1A2
CLASSIFICATION #superfamily cytochrome P450
KEYWORDS electron transfer; endoplasmic reticulum; heme; membrane
protein; monooxygenase; oxidoreductase
FEATURE
456 #binding_site heme iron (Cys) (axial ligand) #status
predicted
SUMMARY #length 513 #molecular-weight 58082 #checksum 3323
Query Match      67.9%; Score 55; DB 5; Length 513;
Best Local Similarity 85.7%; Pred. No. 1.49e+01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 373 lyrytsf 379
||||:|
QY 2 LYRYGSF 8

RESULT 13
ENTRY S44151 #type complete
TITLE cathepsin L (EC 3.4.22.15) - fluke (Schistosoma mansoni)
ORGANISM #formal_name Schistosoma mansoni
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
30-Jun-1995
ACCESSIONS S44151
REFERENCE S44151
#authors Michel, A.; Klinkert, M.; Kunz, W.

```



```
#submission submitted to the EMBL Data Library, April 1994
#description Sequence and expression of cathepsin L from Schistosoma
#accession S44151
#status preliminary
#molecule_type mRNA
#residues 1-317 #label MIC
#cross-references EMBL:232529
KEYWORDS Cysteine proteinase; hydrolase
SUMMARY #length 317 #molecular-weight 36571 #checksum 8871

Query Match 66.7%; Score 54; DB 12; Length 317;
Best Local Similarity 66.7%; Pred. No. 2.13e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 228 lyhygpisv 236
|||:|:|
QY 2 LYRYGSFSV 10

RESULT 14
ENTRY S29129 #type fragment
TITLE calreticulin precursor (clone 3) - African clawed frog
#formal_name Xenopus laevis #common_name African clawed frog
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
03-Mar-1995
ACCESSIONS S29129
REFERENCE S29129
#authors Treves, S.; Zorzato, F.; Pozzan, T.
#journal Biochem. J. (1992) 287:579-581
#title Identification of calreticulin isoforms in the central
nervous system.
#accession S29129
#molecule_type mRNA
#residues 1-411 #label TRE
#cross-references EMBL:X67597
CLASSIFICATION #superfamily calreticulin
KEYWORDS glycoprotein
FEATURE 1-12 #domain signal sequence (fragment) #status predicted
13-411 #label SIG\
408-411 #product calreticulin #status predicted #label MAT\
339 #region endoplasmic reticulum retention signal\
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 411 #checksum 4936

Query Match 66.7%; Score 54; DB 7; Length 411;
Best Local Similarity 66.7%; Pred. No. 2.13e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 300 lysydsfgv 308
||| | | |
QY 2 LYRYGSFSV 10

RESULT 15
ENTRY JC4084 #type complete
TITLE prolly endopeptidase (EC 3.4.-.-) - Pyrococcus furiosus
ALTERNATE_NAMES mlr-2 protein
ORGANISM #formal_name Pyrococcus furiosus
DATE 12-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
21-Mar-1996
ACCESSIONS JC4084
REFERENCE JC4084
#authors Robinson, K.A.; Bartley, D.A.; Robb, F.T.; Schreier, H.J.
#journal Gene (1995) 152:103-106
#title A gene from the hyperthermophile Pyrococcus furiosus whose
deduced product is homologous to members of the prolly
oligopeptidase family of proteases.
#accession JC4084
#molecule_type DNA
```

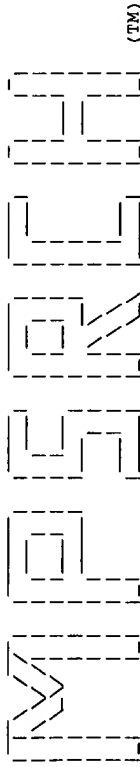
```
##residues 1-616 ##label ROB
##cross-references GB:U08343
KEYWORDS hydrolase; oligopeptidase
FEATURE 477,561,592 #active_site Ser, Asp, His #status predicted
SUMMARY #length 616 #molecular-weight 70867 #checksum 9939

Query Match 66.7%; Score 54; DB 10; Length 616;
Best Local Similarity 60.0%; Pred. No. 2.13e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 335 vllrytsfti 344
||| | | | | |
QY 1 VLRYGSFSV 10

Search completed: Tue Jun 10 11:23:11 1997
Job time : 10 secs.
```

\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:22:33 1997; MasPar time 2.14 Seconds  
Tabular output not generated. 99.074 Million cell updates/sec

Title: >US-08-231-565A-32  
Description: (1-10) from US08231565A.pep  
Perfect Score: 81  
Sequence: 1 VLYRYGFSV 10

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 26.059; Variance 34.980; scale 0.745

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	81	100.0	668	PM17_HUMAN	PMEL 17 PROTEIN PRECU	1.49e+05
2	61	75.3	705	PPCE_FLAME	PROLYL ENDOPEPTIDASE	2.75e-01
3	61	75.3	705	PPCF_FLAME	PROLYL ENDOPEPTIDASE	2.75e-01
4	58	71.6	689	PPCE_AERHY	PROLYL ENDOPEPTIDASE	1.05e+00
5	57	70.4	411	YG4L_YEAST	HYPOTHETICAL 45.2 KD	1.62e+00
6	57	70.4	488	XYNB_CALSA	BETA-XYLIDASE (EC 3	1.62e+00
7	55	67.9	513	CP12_MESAU	CITROCHROME P450 IIA2 (	3.82e+00
8	54	66.7	669	SCRAA_HUMAN	AMILORIDE-SENSITIVE S	5.81e+00
9	53	65.4	212	VF09_VACCV	PROTEIN F9.	8.79e+00
10	53	65.4	212	VF09_VACCC	PROTEIN F9.	8.79e+00
11	53	65.4	212	VF09_VACCV	PROTEIN F9.	8.79e+00
12	52	64.2	774	YVDL_VACCC	HYPOTHETICAL 9.5 KD P	1.32e+01
13	52	64.2	273	YGF0_YEAST	HYPOTHETICAL 30.8 KD	1.32e+01
14	52	64.2	311	DNAL_BACSU	PRIMOSOMAL PROTEIN DN	1.32e+01
15	52	64.2	315	SLTA_BPH19	SHIGA-LIKE TOXIN I SU	1.32e+01
16	52	64.2	315	SLTA_BPH30	SHIGA TOXIN A-CHAIN P	1.32e+01
17	52	64.2	496	11 KYAB_CAEEL	HYPOTHETICAL 55.0 KD	1.32e+01
18	52	64.2	496	8 REPS_STRPY	REPS PROTEIN.	1.32e+01
19	52	64.2	496	8 REPR_STRAG	REPR PROTEIN.	1.32e+01
20	52	64.2	557	TR2M_PSESS	TRYPTOPHAN 2-MONOOXYG	1.32e+01
21	52	64.2	710	7 PPCE_HUMAN	PROLYL ENDOPEPTIDASE	1.32e+01
22	52	64.2	710	7 PPCE_PIG	PROLYL ENDOPEPTIDASE	1.32e+01

RESULT	1	PM17_HUMAN	STANDARD;	PRT;	568 AA.
AC	P40967;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92021023.				
RA	KWON B.-S., CHINTAMANENI C., KOZAK C.A., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,				
RA	KIM K.-K.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).				
CC	!- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; G190106; -				
DR	MIM; 155550; -				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23		POTENTIAL.
FT	CHAIN	24	668		PMEL 17 PROTEIN.
FT	TRANSMEM	575	595		POTENTIAL.
FT	TRANSMEM	603	623		POTENTIAL.
FT	DOMAIN	217	307		POTENTIAL.
FT	DOMAIN	315	444		10 X 13 AA TANDEM REPEATS.
FT	REPEAT	315	327		1.
FT	REPEAT	328	340		2.
FT	REPEAT	341	353		3.
FT	REPEAT	354	366		4.
FT	REPEAT	367	379		5.
FT	REPEAT	380	392		6.
FT	REPEAT	393	405		7.
FT	REPEAT	406	418		8.
FT	REPEAT	419	431		9.
FT	REPEAT	432	444		10.
FT	CARBOHYD	81	81		POTENTIAL.
FT	CARBOHYD	106	106		POTENTIAL.
FT	CARBOHYD	111	111		POTENTIAL.
FT	CARBOHYD	321	321		POTENTIAL.
FT	CARBOHYD	568	568		POTENTIAL.
FT	SEQUENCE	668 AA;	70992 MW;		6E9E1AF0 CRC32;

```

Query Match          100.0%; Score 81; DB 7; Length 668;
Best Local Similarity 100.0%; Pred. No. 1.49e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 476 vlyrygsfsv 485
   |||||
QY 1 VLYRYGSFSV 10

RESULT 2
ID PPCE_FLAME STANDARD; PRT; 705 AA.
AC P27028;
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE PROLYL ENDOPEPTIDASE PRECURSOR (EC 3.4.21.26) (POST-PROLINE CLEAVING
DE ENZYME) (PE) (VERSION 1).
GN FIPEP1.
OS FLAVOBACTERIUM MENINGOSEPTICUM.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC UNCERTAIN.
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE; 92176159.
RA YOSHIMOTO T., KANATANI A., SHIMODA T., INAKA T., KOKUBO T., TSURU D.;
RL J. BIOCHEM. 110:873-878(1991).
RN [2]
SEQUENCE FROM N.A.
RA DIEFENTHAL T.;
RL SUBMITTED (JAN-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CLEAVES PEPTIDE BONDS ON THE C-TERMINAL SIDE OF PROLYL
CC RESIDUES WITHIN PEPTIDES THAT ARE UP TO APPROXIMATELY 30 AMINO
CC ACIDS LONG. HAS AN ABSOLUTE REQUIREMENT FOR AN X-PRO BOND IN THE
CC TRANS CONFIGURATION IMMEDIATELY PRECEDING THE PRO-Y SCISSIBLE
CC BOND.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF PRO-|-XAA >> ALA-|-XAA IN
CC OLIGOPEPTIDES.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9A; ALSO KNOWN AS THE
CC PROLYL OLIGOPEPTIDASE FAMILY.
DR EMBL; X63674; G456523; -
DR PIR; JX0194; JX0194.
DR PROSITE; PS00708; PRO_ENDOPEP_SER.
DR HYDROLASE; SERINE PROTEASE; PERIPLASMIC; SIGNAL.
FT SIGNAL 1 20
FT CHAIN 21 705 PROLYL ENDOPEPTIDASE.
FT ACT_SITE 556 556 CHARGE RELAY SYSTEM.
FT ACT_SITE 675 675 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 110 110 S -> C (IN REF. 2).
FT CONFLICT 587 587 R -> A (IN REF. 2).
SQ SEQUENCE 705 AA; 78707 MW; 714210B4 CRC32;

Query Match          75.3%; Score 61; DB 7; Length 705;
Best Local Similarity 50.0%; Pred. No. 2.75e-01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 472 ilysygsfni 481
   ||| |||
QY 1 VLYRYGSFSV 10

RESULT 3
ID PPCE_FLAME STANDARD; PRT; 705 AA.
AC P27195;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE PROLYL ENDOPEPTIDASE PRECURSOR (EC 3.4.21.26) (POST-PROLINE CLEAVING
DE ENZYME) (PE) (VERSION 2).
OS FLAVOBACTERIUM MENINGOSEPTICUM.

```

```

OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC UNCERTAIN.
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-ATCC 33958;
RX MEDLINE; 92235032.
RA CHEVALLIER S., GOELTZ P., THIBAUT P., BANVILLE D., GAGNON J.;
RL J. BIOL. CHEM. 267:8192-8199(1992).
CC -!- FUNCTION: CLEAVES PEPTIDE BONDS ON THE C-TERMINAL SIDE OF PROLYL
CC RESIDUES WITHIN PEPTIDES THAT ARE UP TO APPROXIMATELY 30 AMINO
CC ACIDS LONG. HAS AN ABSOLUTE REQUIREMENT FOR AN X-PRO BOND IN THE
CC TRANS CONFIGURATION IMMEDIATELY PRECEDING THE PRO-Y SCISSIBLE
CC BOND.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF PRO-|-XAA >> ALA-|-XAA IN
CC OLIGOPEPTIDES.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9A; ALSO KNOWN AS THE
CC PROLYL OLIGOPEPTIDASE FAMILY.
DR EMBL; M81461; G148698; -
DR PROSITE; PS00708; PRO_ENDOPEP_SER.
DR HYDROLASE; SERINE PROTEASE; PERIPLASMIC; SIGNAL.
FT SIGNAL 1 20
FT CHAIN 21 705 PROLYL ENDOPEPTIDASE.
FT ACT_SITE 556 556 CHARGE RELAY SYSTEM.
FT ACT_SITE 675 675 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 705 AA; 78790 MW; EEDC8997 CRC32;

Query Match          75.3%; Score 61; DB 7; Length 705;
Best Local Similarity 50.0%; Pred. No. 2.75e-01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 472 ilysygsfni 481
   ||| |||
QY 1 VLYRYGSFSV 10

RESULT 4
ID PPCE_AERHY STANDARD; PRT; 689 AA.
AC Q06903;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE PROLYL ENDOPEPTIDASE (EC 3.4.21.26) (POST-PROLINE CLEAVING ENZYME)
DE (PE).
OS AEROMONAS HYDROPHILA.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC VIBRONACEAE.
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-AK-9;
RX MEDLINE; 93380910.
RA KANATANI A., YOSHIMOTO T., KITAZONO A., KOKUBO T., TSURU D.;
RL J. BIOCHEM. 113:790-796(1993).
CC -!- FUNCTION: CLEAVES PEPTIDE BONDS ON THE C-TERMINAL SIDE OF PROLYL
CC RESIDUES WITHIN PEPTIDES THAT ARE UP TO APPROXIMATELY 30 AMINO
CC ACIDS LONG. HAS AN ABSOLUTE REQUIREMENT FOR AN X-PRO BOND IN THE
CC TRANS CONFIGURATION IMMEDIATELY PRECEDING THE PRO-Y SCISSIBLE
CC BOND.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF PRO-|-XAA >> ALA-|-XAA IN
CC OLIGOPEPTIDES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9A; ALSO KNOWN AS THE
CC PROLYL OLIGOPEPTIDASE FAMILY.
DR EMBL; D14005; G216201; -
DR PIR; JN0585; JN0585.
DR PROSITE; PS00708; PRO_ENDOPEP_SER.
DR HYDROLASE; SERINE PROTEASE.
FT INIT_MET 0 0
FT ACT_SITE 537 537 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 656 656 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 689 AA; 76384 MW; 8C78BBE2 CRC32;

Query Match          71.6%; Score 58; DB 7; Length 689;

```

```
Best Local Similarity 60.0%; Pred. No. 1.05e+00;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 453 ilyyggfddv 462
:| | | | |
Qy 1 VLRYGFSV 10

RESULT 5
ID YG4L_YEAST STANDARD; PRT; 411 AA.
AC P42942;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 45.2 KD GTP-BINDING PROTEIN IN TRX1-RTAI INTERGENIC
DE REGION.
GN YGR210C OR G7748.
OS SACHAROMYCES CREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA GUERREIRO P., BARREIROS T., SOARES H., CYRNE L., MAIA E SILVA A.,
RA RODRIGUES-POUSADA C.;
RL YEAST 12:273-280(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GRF88;
RA SONG J.M., CHEUNG E., RABINOWITZ J.C.;
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Z49133; G790501; -
DR EMBL; Z72995; E243749; -
DR EMBL; U40843; G1165215; -
KW HYPOTHETICAL PROTEIN; GTP-BINDING.
FT NP_BIND 11 18 GTP (POTENTIAL).
FT NP_BIND 83 87 GTP (POTENTIAL).
SQ SEQUENCE 411 AA; 45214 MW; EC8B5794 CRC32;

Query Match 70.4%; Score 57; DB 11; Length 411;
Best Local Similarity 70.0%; Pred. No. 1.62e+00;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 308 vlyrfgstgv 317
| | | | |
Qy 1 VLRYGFSV 10

RESULT 6
ID XYNB_CALSA STANDARD; PRT; 488 AA.
AC P23552;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE BETA-XYLOSIDASE (EC 3.2.1.37) (1.4-BETA-D-XYLAN XYLOHYDROLASE)
DE (XYLAN 1.4-BETA-XYLOSIDASE).
GN XYNB.
OS CALDOCELLUM SACCHAROLYTICUM
OC PROKARYOTA; NOT YET CLASSIFIED.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90253140.
RA LUETHI E., LOVE D.R., MCANULTY J., WALLACE C., CAUGHEY P.A.,
RA SAUL D., BERGQUIST P.L.;
RL APPL. ENVIRON. MICROBIOL. 56:1017-1024(1990).
CC -!- FUNCTION: BETA-XYLOSIDASE IS AN INTRACELLULAR XYLAN-DEGRADING
CC ENZYME.
CC -!- CATALYTIC ACTIVITY: HYDROLYSES OF 1.4-BETA-D-XYLANS SO AS TO
CC REMOVE SUCCESSIVE D-XYLOSE RESIDUES FROM THE NON-REDUCING
CC TERMINI. IT ALSO HYDROLYSES XYOBILOSE.
CC -!- SIMILARITY: BELONGS TO FAMILY 39 OF GLYCOSYL HYDROLASES.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-7 IS THE INITIATOR.
DR EMBL; M34459; G144300; -
DR PIR; E37202; E37202.
```

```
DR PROSITE: PS01027; GLYCOSYL-HYDROL_F39.
KW XYLAN DEGRADATION; HYDROLASE; GLYCOSIDASE.
FT ACT_SITE 163 PROTON DONOR (POTENTIAL).
FT ACT_SITE 265 NUCLEOPHILE (POTENTIAL).
SQ SEQUENCE 488 AA; 56365 MW; C0926E00 CRC32;

Query Match 70.4%; Score 57; DB 10; Length 488;
Best Local Similarity 44.4%; Pred. No. 1.82e+00;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 181 lykyaafai 189
| | | | |
Qy 2 LYRYGFSV 10

RESULT 7
ID CP12_MESAU STANDARD; PRT; 513 AA.
AC P24453;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CYTOCHROME P450 1A2 (EC 1.14.14.1) (P450-MC4) (METHYLCHOLANTHRENE-
DE INDUCIBLE) (HEPATIC CYTOCHROME P-450MC1).
GN CYP1A2.
OS MESOCRICETUS AURATUS (GOLDEN HAMSTER).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 91112759.
RA LAI T.S., CHIANG J.Y.;
RL ARCH. BIOCHEM. BIOPHYS. 283:429-439(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92138673.
RA SAGAMI I., OHMACHI T., FUJII H., KIKUCHI H., WATANABE M.;
J. BIOCHEM. 110:641-647(1991).
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC ACIDS, AND XENOBIOTICS.
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC OXIDIZED FLAVOPROTEIN + H(2)O.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -!- INDUCTION: BY 3-METHYLCHOLANTHRENE.
CC -!- TISSUE SPECIFICITY: FOUND IN LUNG AND LIVER.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; M63787; G191355; -
DR EMBL; D10252; G220319; -
DR EMBL; D10914; G398133; -
DR PIR; S13885; S13885.
DR PROSITE: PS00086; CYTOCHROME_P450.
KW OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
KW MICROSOME.
FT BINDING 456 456 HEME (BY SIMILARITY).
FT CONFLICT 49 49 I -> F (IN REF. 1).
FT CONFLICT 52 53 HV -> MC (IN REF. 1).
FT CONFLICT 253 254 KN -> GG (IN REF. 1).
FT CONFLICT 326 326 L -> W (IN REF. 1).
FT CONFLICT 356 356 R -> L (IN REF. 1).
FT CONFLICT 485 485 T -> Q (IN REF. 1).
SQ SEQUENCE 513 AA; 58082 MW; 40F0041D CRC32;

Query Match 67.9%; Score 55; DB 2; Length 513;
Best Local Similarity 85.7%; Pred. No. 3.82e+00;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 373 lyrytsf 379
| | | | |
Qy 2 LYRYGSF 8
```

```

RESULT 8
ID SCAA_HUMAN STANDARD; PRT; 669 AA.
AC P37088;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DE AMILORIDE-SENSITIVE SODIUM CHANNEL ALPHA-SUBUNIT (LUNG NA+
DE CHANNEL ALPHA SUBUNIT) (ALPHA HENAC) (NONVOLTAGE-GATED SODIUM CHANNEL
DE 1 ALPHA SUBUNIT) (SCNEA).
GN SCNN1A OR SCNN1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE; 94295729.
RA VOILEY N., LINGUEGLIA E., CHAMPIGNY G., MATTEI M.-G., WALDMANN R.,
RA LAZDUNSKI M., BARBRY P.;
RL PROC. NATL. ACAD. SCI. U.S.A. 91:247-251(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE; 94295729.
RA MCDONALD F.J., SNYDER P.M., MCCRAY P.B., WELSH M.J.;
RL AM. J. PHYSIOL. 266:L728-L734(1994).
CC -!- FUNCTION: SODIUM PERMEABLE NON-VOLTAGE- SENSITIVE ION CHANNEL
CC INHIBITED BY THE DIURETIC AMILORIDE. MEDIATE THE ELECTRODIFFUSION
CC OF THE LUMINAL SODIUM (AND WATER, WHICH FOLLOWS OSMOTICALLY)
CC THROUGH THE APICAL MEMBRANE OF EPITHELIAL CELLS. CONTROLS THE
CC REABSORPTION OF SODIUM IN KIDNEY, COLON, LUNG AND SWEAT GLANDS.
CC ALSO PLAYS A ROLE IN TASTE PERCEPTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SUBUNIT: HETEROTRIMER OF AN ALPHA, BETA AND GAMMA SUBUNIT.
CC -!- DISEASE: DEFECTS IN SCNN1A ARE ONE OF THE CAUSE OF
CC PSEUDOHYPONATREMIA TYPE 1 (PHA1), A RARE SALT WASTING DISEASE
CC CHARACTERIZED BY AN OFTEN FULMINANT PRESENTATION IN THE NEONATAL
CC PERIOD WITH DEHYDRATION, HYPONATREMIA, HYPERKALAEMIA, METABOLIC
CC ACIDOSIS, FAILURE TO THRIVE AND WEIGHT LOSS.
CC -!- SIMILARITY: BELONGS TO THE AMILORIDE-SENSITIVE SODIUM CHANNELS
CC FAMILY.
DR EMBL; X76180; G452650; -.
DR EMBL; L29007; G493605; -.
DR MIM; 600228; -.
DR MIM; 177735; -.
KW IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; GLYCOPROTEIN.
FT DOMAIN 1 85
FT TRANSMEM 86 106 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 107 562 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 563 583
FT DOMAIN 584 669 POTENTIAL.
FT CARBOHYD 232 232 POTENTIAL.
FT CARBOHYD 293 293 POTENTIAL.
FT CARBOHYD 312 312 POTENTIAL.
FT CARBOHYD 397 397 POTENTIAL.
FT CARBOHYD 511 511 POTENTIAL.
SQ SEQUENCE 669 AA; 75703 MW; 2E0E9EA6 CRC32;

Query Match 66.7%; Score 54; DB 9; Length 669;
Best Local Similarity 62.5%; Pred. No. 5.81e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 161 lykysft 168
QY 2 LYRYGSES 9

RESULT 9
ID VF09_VARV STANDARD; PRT; 212 AA.
AC P33869;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)

```

```

DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE PROTEIN F9.
GN F9L OR C13L.
OS VARIOIA VIRUS.
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIA-1967 / ISOLATE IND3;
RX MEDLINE; 94152154.
RA SHELKUNOV S.N., BLINOV V.M., RESENCHUK S.M., TOTMENIN A.V.,
RA SANDAKHCHIEV L.S.;
RL VIRUS RES. 30:239-258(1993).
RN [2]
RP COMPLETE GENOME.
RC STRAIN=INDIA-1967 / ISOLATE IND3;
RX MEDLINE; 93202281.
RA SHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;
RL FEBS LETT. 319:80-83(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BANGLADESH-1975;
RX MEDLINE; 94088747.
RA MASSUNG R.F., ESPOSITO J.J., LIU L., QI J., UTTERBACK T.R.,
RA KNIGHT J.C., AUBIN L., YURAN T.E., PARSONS J.M., LOPAREV V.N.,
RA SELIVANOV N.A., CAVALLARO K.F., KERLAVAGE A.R., MAHY B.W.J.,
RA VENTER C.J.;
RL NATURE 366:748-751(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SOMALIA-1977, CONGO-1965, AND GARCIA-1966;
RA MASSING R.F., LOPAREV V.N., KNIGHT J.C., CHIZHIKOV V.E., PARSONS J.M.,
RA TOTMENIN A.V., SHELKUNOV S.N., ESPOSITO J.J.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; X69198; G297214; -.
DR EMBL; L22579; G438951; -.
DR EMBL; U18340; G885831; -.
DR EMBL; U18337; G885722; -.
DR EMBL; U18338; G885764; -.
DR PIR; D36840; D36840.
SQ SEQUENCE 212 AA; 23746 MW; 32BBA02F CRC32;

Query Match 65.4%; Score 53; DB 10; Length 212;
Best Local Similarity 83.3%; Pred. No. 8.79e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 204 yrygtf 209
QY 3 YRYGSF 8

RESULT 10
ID VF09_VACCC STANDARD; PRT; 212 AA.
AC P21018;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE PROTEIN F9.
GN F9L OR F9.
OS VACCINIA VIRUS (STRAIN COPENHAGEN), AND VACCINIA VIRUS (STRAIN L-IVP).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COPENHAGEN;
RX MEDLINE; 91021027.
RA GOEBEL S.-J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RL VIROLOGY 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RC STRAIN=COPENHAGEN;
RA GOEBEL S.-J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,

```

```
RA PAOLETTI E.;
RL VIROLOGY 179:517-563(1990).
RN [3]
RC SEQUENCE FROM N.A.
RC STRAIN=L-IVP;
RA MIKRYUKOV N.N., CHIZHIKOV V.E., PRIKHOD'KO G.G., URMANOV I.M.,
RA SERPINSKII O.I., BLINOV V.M., NIKULIN A.E., VASILENKO S.K.;
RL BIOTEKHOLOGIYA 4:442-449(1988).
DR EMBL; M35027; G335373; -
DR EMBL; M57977; G335705; -
DR PIR; E42507; E42507.
FT SIMILAR 20 160
FT 31.3% IDENTITY TO FP2 PROTEIN OF
FT FOWLPOX VIRUS (AC F15910).
SQ SEQUENCE 212 AA; 23792 MW; 4579EFA4 CRC32;

Query Match 65.4%; Score 53; DB 10; Length 212;
Best Local Similarity 83.3%; Pred. No. 8.79e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 204 yrygtf 209
QY 3 YRYGSF 8

RESULT 11
ID VF09_VACCV STANDARD; PRT; 212 AA.
AC P24361;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE PROTEIN F9.
GN F9L.
OS VACCINIA VIRUS (STRAIN WR).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91020979.
RA ROSEMAN N.A., SLABAUGH M.B.;
RL VIROLOGY 178:410-418(1990).
DR EMBL; M34368; G335619; -
DR PIR; A36213; A36213.
FT SIMILAR 20 160
FT 31.3% IDENTITY TO FP2 PROTEIN OF
FT FOWLPOX VIRUS (AC F15910).
SQ SEQUENCE 212 AA; 23962 MW; E44CB668 CRC32;

Query Match 65.4%; Score 53; DB 10; Length 212;
Best Local Similarity 83.3%; Pred. No. 8.79e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 204 yrygtf 209
QY 3 YRYGSF 8

RESULT 12
ID YVDI_VACCC STANDARD; PRT; 74 AA.
AC P20553;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 9.5 KD PROTEIN.
GN D ORF 1.
OS VACCINIA VIRUS (STRAIN COPENHAGEN).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91021027.
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RL VIROLOGY 179:247-266(1990).
RN [2]
```

```
RP COMPLETE GENOME.
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RL VIROLOGY 179:517-563(1990).
DR EMBL; M35027; G335463; -
DR PIR; B42517; B42517.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 74 AA; 9485 MW; F0E76B25 CRC32;

Query Match 64.2%; Score 52; DB 11; Length 74;
Best Local Similarity 83.3%; Pred. No. 1.32e+01;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 69 yrygnf 74
QY 3 YRYGSF 8

RESULT 13
ID YGF0_YEAST STANDARD; PRT; 273 AA.
AC P53177;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 30.8 KD PROTEIN IN OLE1-TIF4632 INTERGENIC REGION.
GN YGL050W.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RA FEUERMAN M., POTIER S., SOUCIET J.L.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Z72572; E243282; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 273 AA; 30806 MW; 89427531 CRC32;

Query Match 64.2%; Score 52; DB 11; Length 273;
Best Local Similarity 50.0%; Pred. No. 1.32e+01;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 127 llykyepf 134
QY 1 VLRYGSF 8

RESULT 14
ID DNAL_BACSU STANDARD; PRT; 311 AA.
AC P06567;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE PRIMOSOMAL PROTEIN DNAL.
GN DNAL.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87117549.
RA OGASAWARA N., MORIYA S., MAZZA P.G., YOSHIKAWA H.;
RL NUCLEIC ACIDS RES. 14:9989-9999(1986).
RN [2]
RP SEQUENCE OF 1-206 FROM N.A.
RX MEDLINE; 87118226.
RA HOSHINO T., MCKENZIE T., SCHMIDT S., TANAKA T., SUEOKA N.;
RL PROC. NATL. ACAD. SCI. U.S.A. 84:653-657(1987).
RN [3]
RP SIMILARITY TO DNAA.
RX MEDLINE; 92195821.
RA KOONIN E.V.;
RL NUCLEIC ACIDS RES. 20:1143-1143(1992).
RN [4]
RP IDENTIFICATION.
RX MEDLINE; 95291463.
```

Search completed: Tue Jun 10 11:22:43 1997  
Job time : 10 secs.

```
RA BRUAND C., EHRLICH S.D.;
RL MICROBIOLOGY 141:1199-1200(1995).
CC -!- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -!- SIMILARITY: SOME, TO DNAA FROM VARIOUS BACTERIA.
DR EMBL; X04963; G39881; -.
DR EMBL; M15183; G468268; -.
DR PIR; B24720; I0BS44.
DR PIR; C26580; C26580.
DR SUBTILIST; BGI0359; DNAL.
KW PRIMOSOME; DNA REPLICATION; ATP-BINDING.
FT NP_BIND 168 175 ATP (PROBABLE).
FT CONFLICT 19 19 K -> N (IN REF. 2).
FT CONFLICT 24 24 M -> T (IN REF. 2).
SQ SEQUENCE 311 AA; 36114 MW; 61C67D01 CRC32;

Query Match 64.2%; Score 52; DB 3; Length 311;
Best Local Similarity 66.7%; Pred. No. 1.32e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 164 lylygkfgv 172
   |||||
Qy 2 LYRYGSFSV 10

RESULT 15
ID SLTA_BPH19 STANDARD; PRT; 315 AA.
AC P08026;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE SHIGA-LIKE TOXIN I SUBUNIT A PRECURSOR (VEROTOXIN 1 SUBUNIT A) (SLT-I) (RNA N-GLYCOSIDASE) (EC 3.2.2.22).
GN SLTA.
OS BACTERIOPHAGE H19B.
OC VIRIDAE; NOT YET CLASSIFIED.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87260808.
RA CALDERWOOD S.B., AUCLAIR F., DONOHUE-ROLFE A., KEUSCH G.T.,
RA MEKALANOS J.J.;
RL PROC. NATL. ACAD. SCI. U.S.A. 84:4364-4368(1987).
RN [2]
RP ACTIVE SITE.
RX MEDLINE; 88190113.
RA HOYDE C.J., CALDERWOOD S.B., MEKALANOS J.J., COLLIER R.J.;
RL PROC. NATL. ACAD. SCI. U.S.A. 85:2568-2572(1988).
CC -!- FUNCTION: THE SUBUNIT A IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE SPECIFIC ADENOSINE ON THE 28S RNA.
CC -!- SUBUNIT: THE SHIGA-LIKE TOXINS CONTAIN A SINGLE A SUBUNIT AND MULTIPLE COPIES OF A B SUBUNIT.
CC -!- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING PROTEINS.
DR EMBL; M16625; G215044; -.
DR PIR; A27052; XUBPH9.
DR PROSITE; PS00275; SHIGA_RICIN.
KW HYDROLASE; GLYCOSIDASE; TOXIN; SIGNAL.
FT SIGNAL 1 22
FT CHAIN 23 315 SHIGA-LIKE TOXIN I SUBUNIT A.
FT ACT_SITE 189 189
SQ SEQUENCE 315 AA; 34799 MW; 671054B3 CRC32;

Query Match 64.2%; Score 52; DB 9; Length 315;
Best Local Similarity 55.6%; Pred. No. 1.32e+01;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 110 vfyrfadfs 118
   ||::||
Qy 1 VLYRYGSFS 9
```

\*\*\*\*\*  
WQSEH (TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:24:50 1997; MasPar time 1.89 Seconds  
Tabular output not generated. 57.570 Million cell updates/sec

Title: >US-08-231-565A-33  
Description: (1-10) from US08231565A.pep  
Perfect Score: 72  
Sequence: 1 ALDGGNKHFL 10

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 16.646; Variance 42.058; scale 0.396

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	72	100.0	10 15	R82157	Melanoma-specific mut	1.92e-02
2	72	100.0	10 15	R84201	gp100 melanoma antige	1.92e-02
3	72	100.0	661 14	R78646	Melanoma associated a	1.92e-02
4	72	100.0	661 15	R84855	MART-1 melanoma antig	1.92e-02
5	72	100.0	661 15	R84854	MART-1 melanoma antig	1.92e-02
6	56	77.8	147 14	R80521	Rat allograft inflam	2.82e+00
7	48	66.7	410 11	R60618	Murine type II IL-1 r	2.97e+01
8	48	66.7	410 3	R15864	Murine type II interl	2.97e+01
9	48	66.7	410 15	R85481	Mouse type II interle	2.97e+01
10	47	65.3	846 14	R74692	Rat very low density	3.96e+01
11	47	65.3	873 8	R44735	apo-E lipoprotein rec	3.96e+01
12	47	65.3	936 9	R52579	Recombinant collagena	3.96e+01
13	45	62.5	195 10	R57024	Partial sequence of h	6.96e+01
14	45	62.5	249 1	P81193	Biotin biosynthesis g	6.96e+01
15	45	62.5	353 10	R57011	Human A transferase.	6.96e+01
16	45	62.5	353 2	R11317	Histo-blood gp. A gly	6.96e+01
17	45	62.5	354 10	R57016	Human B transferase (	6.96e+01
18	45	62.5	354 10	R57013	Human A transferase (	6.96e+01
19	45	62.5	354 10	R57020	Human A transferase (	6.96e+01
20	45	62.5	354 2	R11792	Histo-blood gp. B gly	6.96e+01

21	45	62.5	354 2	R11790	Histo-blood gp. B gly	6.96e+01
22	45	62.5	354 2	R11789	Histo-blood gp. A gly	6.96e+01
23	45	62.5	358 10	R57014	Human B transferase (	6.96e+01
24	45	62.5	375 10	R57021	Human A transferase (	6.96e+01
25	45	62.5	516 3	P61362	Soybean glycinin A3p4	6.96e+01
26	45	62.5	561 3	P61363	Soybean glycinin A5p4	6.96e+01
27	45	62.5	846 14	R74691	Human very low densit	6.96e+01
28	45	62.5	873 19	W02212	Human VLDL receptor.	6.96e+01
29	44	61.1	816 13	R66931	AMML chromosome inv(1	9.20e+01
30	44	61.1	864 17	R93021	Human glucagon degrad	9.20e+01
31	44	61.1	885 13	R66930	AMML chromosome inv(1	9.20e+01
32	44	61.1	1474 2	R11334	Recombinant human alp	9.20e+01
33	44	61.1	1484 2	R11749	Human alpha-2 macrogli	9.20e+01
34	43	59.7	526 1	P80986	Sequence of fusion pr	1.21e+02
35	43	59.7	526 4	R20501	Newcastle Disease Vir	1.21e+02
36	43	59.7	553 11	R49141	Newcastle disease vir	1.21e+02
37	43	59.7	553 2	P70176	Sequence of Newcastle	1.21e+02
38	43	59.7	553 11	R58858	Newcastle Disease Vir	1.21e+02
39	43	59.7	553 3	R14480	Newcastle disease vir	1.21e+02
40	43	59.7	553 2	R10065	Newcastle disease vir	1.21e+02
41	43	59.7	3144 11	R58777	Protein encoded by Hu	1.21e+02
42	42	58.3	496 1	P82687	Beta-amylose from pla	1.59e+02
43	42	58.3	535 19	R91310	Fungal signal recogni	1.59e+02
44	42	58.3	1169 18	R96126	Bacillus thuringiensis	1.59e+02
45	42	58.3	4544 9	R47861	Alpha 2-Macroglobulin	1.59e+02

ALIGNMENTS

RESULT 1  
ID R82157 standard; peptide; 10 AA.  
AC R82157;  
DT 25-MAR-1996 (first entry)  
DE Melanoma-specific mutant immunogen epitope 10mer peptide.  
KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
KW Cytotoxic T cell; lymphocyte; HLA-A2.  
OS Homo sapiens.  
PN W09523561-A2.  
PD 24-AUG-1995.  
PF 16-FEB-1995; U01991.  
PR 16-FEB-1994; US-197399.  
PR 29-APR-1994; US-234784.  
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.  
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
DR WPI: 95-302688/39  
PT Melanoma-specific immunogen comprises epitope(s) homologous with  
PT pMEL-17 - are highly potent stimulators of HLA-A2+CTL's useful in  
PT adoptive immuno-therapy  
PS Example 8; Page 52; 148pp; English.  
CC A melanoma-specific immunogen homologous with pMEL-17 comprises one  
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
CC R82108 are of particular interest. The immunogen can be used for  
CC partial protection in mammals against melanoma peptides which are  
CC homologous with pMEL-17 are highly potent stimulators of HLA-A2+  
CC CTLs in several cell lines and can be used in immunotherapy or  
CC incorporated into immunogenic conjugates as vaccines.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 72; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred.No. 1.92e-02;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 alidggnkhfl 10  
QY 1 ALDGGNKHFL 10

RESULT 2  
ID R84201 standard; Peptide; 10 AA.  
AC R84201;  
DT 25-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G10-22).



KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA,  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 55; Page 131; 184pp; English.  
 CC The immunogenic peptide is derived from cDNA25 (R84854), a  
 CC melanoma antigen derivative of gp100 (see R84855). The  
 CC peptide and its derivatives (see R84200-R84211) are used in  
 CC medicaments (vaccines) for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against melanoma-specific  
 CC antigens and its immunogenic peptides may be used in the  
 CC detection and isolation of the antigen from a sample, the  
 CC detection of which is indicative of a disease state  
 CC (melanoma or metastatic melanoma).  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 72; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.92e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 aldggnkhfl 10

Qy 1 ALDGGNKHFL 10

# RESULT

ID R78646 standard; Protein; 661 AA.  
 AC R78646;  
 DT 22-JAN-1996 (first entry)  
 DE Melanoma associated antigen gp100.  
 KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 OS Homo sapiens.  
 PN EP-668350-A1.  
 PD 23-AUG-1995.  
 PF 14-FEB-1995; 200348.  
 PR 16-FEB-1994; EP-200337.  
 PR 21-DEC-1994; EP-203709.  
 PA (ALKU ) AKZO NOBEL NV.  
 PI Adema GJ, Figdor CG,  
 DR WPI; 95-284790/38.  
 DR N-PSDB; Q96055.  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PT detection of tumours  
 PS Claim 1; Page 22-24; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC may be used in the production of vaccines. Nucleotide sequences  
 CC encoding the immunogenic peptides may be used as primers and probes  
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
 CC capable of binding to the melanoma associated antigen can be  
 CC cultured ex vivo and returned to melanoma particles, and when  
 CC radiolabelled, they may be used to identify tumour deposits.  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 72; DB 14; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 1.92e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 aldggnkhfl 233

Qy 1 ALDGGNKHFL 10

## RESULT

ID R84855 standard; Protein; 661 AA.  
 AC R84855;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen gp100.  
 KW gp100; MART-1; melanoma antigen recognised by T-cell;  
 KW cDNA25 antigen derivative; melanocyte; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian sp.

Key Location/Qualifiers

FT Peptide 154..163  
 FT /label= G9-154-immunogenic\_peptide  
 FT Peptide 208..217  
 FT /label= G9-209-immunogenic\_peptide  
 FT /note= "see R84210"  
 FT Peptide 280..288  
 FT /label= G9-280-immunogenic\_peptide  
 FT /note= "see R84208"  
 FT Peptide 457..266  
 FT /label= immunogenic\_peptide  
 FT Peptide 476..485  
 FT /label= immunogenic\_peptide  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA,  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 7A; 184pp; English.  
 CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which  
 CC is recognized by T-lymphocytes, and is a derivative of the  
 CC melanoma-specific antigen cDNA25 (see R84854). gp100 is a source  
 CC of immunogenic peptides which are optionally modified to enhance  
 CC their binding to a MHC molecule, and used in medicaments,  
 CC especially vaccines, for the treatment or prevention (by  
 CC immunisation) of melanoma. Antibodies against cDNA2 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 72; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 1.92e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 aldggnkhfl 233

Qy 1 ALDGGNKHFL 10

## RESULT

ID R84854 standard; Protein; 661 AA.  
 AC R84854;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen cDNA25.  
 KW cDNA25; MART-1; melanoma antigen recognised by T-cell;  
 KW gp100 antigen derivative; melanoma; metastatic melanoma;  
 KW tumour-associated antigen; immunogen; diagnosis; prognosis;  
 KW prophylaxis; therapy; vaccine.  
 OS Mammalian sp.

Key Location/Qualifiers

FT Peptide 457..466  
 FT /label= antigenic\_peptide

```

FT /note= "see R84199"
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SPC DEPT HEALTH.
PI Kawakami I, Rosenberg SA;
DR WPI; 95-382963/49.
DR N-PSDB; T02716.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Claim 81; Fig 5A; 184pp; English.
CC CDNA2 is a melanoma antigen (MART-1) which is recognized by
CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-
CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source
CC of immunogenic peptides (see R84199) which are optionally modified
CC (see R84200-R84211) to enhance their binding to a MHC molecule and
CC used in medicaments, especially vaccines, for the treatment or
CC prevention (by immunisation) of melanoma. Antibodies against cDNA2
CC and its immunogenic peptides may be used in the detection and
CC isolation of the antigen from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 661 AA;

Query Match 100.0%; Score 72; DB 15; Length 661;
Best Local Similarity 100.0%; Pred. No. 1.92e-02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 aldggnkhfl 233
   |||||
QY 1 ALDGGNKHFL 10

RESULT 6
ID R80521 standard; Protein; 147 AA.
AC R80521;
DT 08-FEB-1996 (first entry)
DE Rat allograft inflammatory factor 1 (AIF-1).
KW AIF-1; allograft inflammatory factor 1; transplant rejection;
KW inhibitor; immunogenic; detection; diagnosis.
OS Rattus sp.
PN WO9517506-A1.
PD 29-JUN-1995.
PF 21-DEC-1994; U14724.
PR 21-DEC-1993; US-171385.
PA (HARD ) HARVARD COLLEGE.
PI Russell ME, Utans U;
DR WPI; 95-240668/31.
DR N-PSDB; Q99370.
PT DNA encoding allograft rejection factors and immunogenic fragments -
PT useful for identifying transplant rejection inhibitors
PS Claim 8; Page 88; 138pp; English.
CC R80521 is the rat allograft inflammatory factor, AIF-1. The AIF-1
CC gene is a differentially expressed allograft gene which is expressed
CC in allograft tissue during transplant rejection. Identification of
CC the rat AIF-1 product or transcript indicates that allograft
CC rejection is taking place. The rat AIF-1 gene and product are
CC therefore useful in the diagnosis of transplant rejection.
CC The diagnostic methods used allow rejection (vascular inflammation)
CC to be detected at an early stage and require only a small amount of
CC biopsy material.
SQ Sequence 147 AA;

Query Match 77.8%; Score 56; DB 14; Length 147;
Best Local Similarity 88.9%; Pred. No. 2.82e-00;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 24 ldggnkhfl 32
   |||
QY 2 LDGGNKHFL 10

/note= "see R84199"
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SPC DEPT HEALTH.
PI Kawakami I, Rosenberg SA;
DR WPI; 95-382963/49.
DR N-PSDB; T02716.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Claim 81; Fig 5A; 184pp; English.
CC CDNA2 is a melanoma antigen (MART-1) which is recognized by
CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-
CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source
CC of immunogenic peptides (see R84199) which are optionally modified
CC (see R84200-R84211) to enhance their binding to a MHC molecule and
CC used in medicaments, especially vaccines, for the treatment or
CC prevention (by immunisation) of melanoma. Antibodies against cDNA2
CC and its immunogenic peptides may be used in the detection and
CC isolation of the antigen from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 661 AA;

Query Match 100.0%; Score 72; DB 15; Length 661;
Best Local Similarity 100.0%; Pred. No. 1.92e-02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 aldggnkhfl 233
   |||||
QY 1 ALDGGNKHFL 10

RESULT 7
ID R60618 standard; Protein; 410 AA.
AC R60618;
DT 10-MAY-1995 (first entry)
DE Murine type II IL-1 receptor
DE Murine type II IL-1 receptor
KW Rheumatoid arthritis; diabetes; multiple sclerosis.
OS Mus musculus.
FH Key Location/Qualifiers
FT Peptide 1..13
FT /label= sig_peptide
PN US5350683-A.
PD 27-SEP-1994.
PR 05-JUN-1990; 534193.
PR 05-JUN-1990; US-534193.
PR 24-AUG-1990; US-573576.
PR 13-DEC-1990; US-627071.
PR 16-MAY-1991; US-701415.
PR 12-JUL-1993; US-091519.
PA (IMMU ) IMMUNEX CORP.
PI Cosman DJ, Dower SK, Lupton SD, Mosley BA, Sims JE;
DR WPI; 94-340288/42.
DR N-PSDB; Q71560.
PT DNA encoding type II interleukin-1 receptors - which can be used
PT in therapeutic compns. to suppress IL-1 dependent immune
PT responses in humans.
PS Claim 1; Columns 41-44; 31pp; English.
CC Q71560 encodes R60618 the murine type II interleukin-1 (IL-1)
CC receptor isolated from the 702/3 cell line. This protein along with a
CC suitable diluent and carrier forms a therapeutic composition, which can
CC be used to suppress IL-1 dependent immune responses in humans. The
CC soluble protein can also be used in the clinical treatment of
CC autoimmune diseases, such as rheumatoid arthritis, diabetes and
CC multiple sclerosis.
SQ Sequence 410 AA;

Query Match 66.7%; Score 48; DB 11; Length 410;
Best Local Similarity 77.8%; Pred. No. 2.97e+01;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 188 ldkgnkefl 196
   |||||
QY 2 LDGGNKHFL 10

RESULT 8
ID R15864 standard; Protein; 410 AA.
AC R15864;
DT 16-MAR-1992 (first entry)
DE Murine type II interleukin-1 receptor.
KW IL-1R; cytokine receptor.
OS Mus musculus.
FH Key Location/Qualifiers
FT Peptide 1..13
FT /label= signal
FT Protein 14..410
FT /note= "IL-1 receptor"
FT Region 356..381
FT /label= transmembrane
PN EP-460846-A.
PD 11-DEC-1991.
PF 24-MAY-1991; 304755.
PR 05-JUN-1990; US-534193.
PR 24-AUG-1990; US-573576.
PR 13-DEC-1990; US-627071.
PA (IMMU-) IMMUNEX CORP.
PI Sims J, Cosman DJ, Lupton SD, Mosley B, Dower SK;
DR WPI; 91-363310/50.
DR N-PSDB; Q15268.
PT Deoxyribonucleic acid encoding interleukin-1 receptor protein -
PT useful as immuno-suppressive and antiinflammatory e.g. in graft
PT rejection and auto-immune disease

```

```

Query Match          65.3%;   Score 47;   DB 8;   Length 873;
Best Local Similarity 50.0%;   Pred. NO. 3.96e+01;
Matches 5;   Conservative 4;   Mismatches 1;   Indels 0;   Gaps 0;

Db      541   tldgakrkf1 550
      :|||::: ||
QY      1   ALDGGNKHFL 10

RESULT 12
ID    p52570 standard. Protein. 936 aa

```

R52579;  
 AC 18-JUN-1994 (first entry)  
 DE Recombinant collagenase.  
 KW Collagenase; Clostridium histolyticum; recombinant; high mol.wt.;  
 KW Escherichia coli; tissue digestion; tumor disassociation; slipped;  
 KW disc; pancreas islet dispersal.  
 OS Clostridium histolyticum.  
 PN W09400580-A.  
 PD 06-JAN-1994.  
 PF 22-JUN-1993; U05944.  
 PR 22-JUN-1992; US-902129.  
 PA (TRIG-) TRIGEN INC.  
 PI Lei S, Lin H;  
 DR WPI; 94-026216/03.  
 DR P-PSDB; Q54844.  
 PT New recombinant DNA encoding high mol.wt. Clostridium  
 PT histolyticum collagenase - e.g. for releasing cells from tissue  
 PS or treating slipped discs  
 PS Disclosure; Page 36-41; 64pp; English.  
 CC A collagenase gene was identified in a pRK290 library containing C.  
 CC histolyticum ATCC 21000 DNA. The insert was placed in opposite  
 CC orientations in pC16 and pC17, serially deleted and sequenced  
 CC (Q54844). The gene was incorporated into a plasmid vector for  
 CC expression in a host cell, esp. Escherichia coli. Recombinant  
 CC collagenase (R52579) is 25-100% more active than native collagenases,  
 CC and can be used for any application in which tissue digestion is  
 CC required. Termination signal Q54845 is located 11 bp downstream of  
 CC the translation terminator, and consists of a stem-loop structure  
 CC followed by an AT-rich area.  
 SQ Sequence 936 AA;  
 Query Match 65.3%; Score 47; DB 9; Length 936;  
 Best Local Similarity 55.6%; Pred. No. 3.96e+01;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 221 lddaekhy1 229  
 QY 2 LDGGRKHF 10  
 RESULT 13  
 ID R57024 standard; Protein; 195 AA.  
 AC R57024;  
 DT 16-FEB-1995 (first entry)  
 DE Partial sequence of human A transferase.  
 KW Blood; group; determinant; antigen; erythrocyte; oligosaccharide;  
 KW glycoconjugate; glycosphingolipid; glycoprotein; glycosyltransferase;  
 KW transferase.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Region  
 FT 3  
 FT /note= "Site of deletion."  
 FN US3326857-A.  
 PD 05-JUL-1994.  
 PF 31-AUG-1989; 402695.  
 PR 31-AUG-1989; US-402695.  
 PR 29-AUG-1991; US-752101.  
 PA (BIOM-) BIOMEMBRANE INST.  
 PI Clausen H, Hakomori S, White T, Yamamoto F;  
 DR WPI; 94-217098/26.  
 DR N-PSDB; Q68826.  
 PT Isolated DNA molecules - encode human histo-blood groups A-, B-  
 PT and O-glycotransferases  
 PS Example 9; Figure 10; 63pp; English.  
 CC The histo-blood group ABH determinants are major allogeneic antigens  
 CC in both erythrocytes and tissues of humans. They generally  
 CC constitute peripheral parts of the oligosaccharide chains of  
 CC glycoconjugates i.e. linked to lipids (glycosphingolipids) or to  
 CC proteins (glycoproteins). It was proposed that the A and B  
 CC phenotypes were associated with glycosyltransferases that converted  
 CC the H substance associated with the O phenotype to A and B  
 CC respectively, through the addition of alpha1-3-N-acetylgalactosamine  
 CC or alpha1-3-galactosyl residues to the H antigen Fuc-alpha1-2Gal-

CC betal-R. Hence, the primary products of the histo-blood group A  
 CC and B genes are the respective glycosyltransferases. The full  
 CC sequence of the human transferase A is described in R57011.  
 SQ Sequence 195 AA;  
 Query Match 62.5%; Score 45; DB 10; Length 195;  
 Best Local Similarity 44.4%; Pred. No. 6.96e+01;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Db 9 letaekhm 17  
 QY 2 LDGGRKHF 10  
 RESULT 14  
 ID P81193 standard; protein; 249 AA.  
 AC P81193;  
 DT 04-DEC-1990 (first entry)  
 DE Biotin biosynthesis gene fragment #2 of B.sphaericus IFO 3525.  
 KW biotin biosynthesis; fermentative biotin production;  
 KW plasmid pTG1418;  
 KW bio C; bio F; bio H; seborrhagic dermatitis treatment; ss.  
 OS Bacillus sphaericus.  
 PN EP-266240-A.  
 PD 04-MAY-1988.  
 PF 28-SEP-1987; 402157.  
 PR 30-SEP-1986; FR-013603.  
 PR 18-MAY-1987; FR-006916.  
 PA (TRAN-) Transgene SA.  
 PI Gloekler R, Speck D, Lemoine Y;  
 DR WPI; 88-121306/18.  
 DR N-PSDB; N81551.  
 PT New DNA sequences encoding enzymes involved in biotin biosynthesis -  
 PT isolated from Bacillus sphaericus, and transformants useful in  
 PT fermentative biotin production.  
 PS Disclosure; pp: French.  
 CC DNA was isolated from B.sphaericus IFO 3525, cut with HindIII and  
 CC the fragments sub-cloned into pBR322. The recombinant plasmids were  
 CC used to transform E.coli bio- mutants. Plasmid pTG1418 was present  
 CC in one clone (E.coli C600 CNCM I-609) which showed complementation  
 CC for bio C and bio F. The plasmid contains a 4.53kb insert  
 CC which includes 3 long open reading frames (LORFs) as described in the  
 CC Features table.  
 CC A second clone was isolated which contained DNA encoding bio D, bio  
 CC B and bio A. The insert present in the plasmid of this clone and the  
 CC 4.3kb insert from pTG1400 were cloned together in pBR322 to produce a  
 CC plasmid (pTG1440) which complements biotin auxotrophy in a bio delta  
 CC FCD mutant.  
 CC Transformed cells are used to produce biotin by culture in a medium  
 CC contg pimelic acid and/or biotin vitamer. Co-culture of two  
 CC transformants is preferred where one synthesises the vitamer from  
 CC pimelic acid and the other converts the vitamer to biotin.  
 CC See also N81546.  
 SQ Sequence 249 AA;  
 Query Match 62.5%; Score 45; DB 1; Length 249;  
 Best Local Similarity 62.5%; Pred. No. 6.96e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Db 21 leggekhi 28  
 QY 2 LDGGRKHF 9  
 RESULT 15  
 ID R57011 standard; Protein; 353 AA.  
 AC R57011;  
 DT 15-FEB-1995 (first entry)  
 DE Human A transferase.  
 KW Blood; group; determinant; antigen; erythrocyte; oligosaccharide;  
 KW glycoconjugate; glycosphingolipid; glycoprotein; glycosyltransferase;  
 KW transferase.  
 OS Homo sapiens.

PN US5326857-A.  
 PD 05-JUL-1994.  
 PF 31-AUG-1989; 402695.  
 PR 31-AUG-1989; US-402695.  
 PR 29-AUG-1991; US-752101.  
 PA (BIOM-) BIOMEMBRANE INST.  
 PI Clausen H, Hakomori S, White T, Yamamoto F;  
 DR WPI; 94-217098/26.  
 DR N-PSDB; Q68806.  
 PT Isolated DNA molecules - encode human histo-blood groups A-, B-  
 and O-glycotransferases  
 PS Disclosure; Figure 3; 63pp; English.  
 CC The histo-blood group ABH determinants are major allogeneic antigens  
 in both erythrocytes and tissues of humans. They generally  
 constitute peripheral parts of the oligosaccharide chains of  
 glycoconjugates i.e. linked to lipids (glycosphingolipids) or to  
 proteins (glycoproteins). It was proposed that the A and B  
 phenotypes were associated with glycosyltransferases that converted  
 the H substance associated with the O phenotype to A and B  
 respectively, through the addition of alpha1-3-N-acetylgalactosamine  
 or alpha1-3-galactosyl residues to the H antigen Fuc-alpha1-2Gal-  
 beta1-R. Hence, the primary products of the histo-blood group A  
 and B genes are the respective glycosyltransferases.  
 SQ Sequence 353 AA;

Query Match 62.5%; Score 45; DB 10; Length 353;  
 Best Local Similarity 44.4%; Pred. No. 6.96e+01;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 133 letaekhf 141  
 QY 2 LDGKNKHF 10

Search completed: Tue Jun 10 11:24:59 1997  
 Job time : 9 secs.



```

ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS    30-Sep-1993
REFERENCE     A41234
#authors      Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
              Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
              Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
              coat color locus on mouse chromosome 10 and is in a
              syntenic region on human chromosome 12.
#cross-references MIM:92021023
#accession    A41234
#status       preliminary
#molecule_type mRNA
#residues     1-668 #label KWO
#cross-references GB:W77348
SUMMARY       #length 668 #molecular-weight 70932 #checksum 6409
Query Match   100.0%; Score 72; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 3.17e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 aldggnkhfl 233
|||||
QY 1 ALDGGNKHFL 10

RESULT 3
ENTRY   S53871 #type complete
TITLE   Pmel 17 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE    27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
ACCESSIONS S53871
REFERENCE   S53871
#authors    Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
              Chintamaneni, C.; Bennett, D.; Pickard, R.T.
#journal    Nucleic Acids Res. (1995) 23:154-158
#title      Mouse silver mutation is caused by a single base insertion in
              the putative cytoplasmic domain of Pmel 17.
#accession  S53871
#status     preliminary
#molecule_type mRNA
#residues   1-626 #label KWO
SUMMARY     #length 626 #molecular-weight 65979 #checksum 5710
Query Match   81.9%; Score 59; DB 14; Length 626;
Best Local Similarity 80.0%; Pred. No. 3.96e-02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 224 aldggetkhfl 233
|||||
QY 1 ALDGGNKHFL 10

RESULT 4
ENTRY   I55617 #type complete
TITLE   allograft inflammatory factor-1 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE    26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
ACCESSIONS I55617
REFERENCE   I55617
#authors    Utans, U.; Arceci, R.J.; Yamashita, Y.; Russell, M.E.
#journal    J. Clin. Invest. (1995) 95:2954-2962
#title      Cloning and characterization of allograft inflammatory
              factor-1: a novel macrophage factor identified in rat
              cardiac allografts with chronic rejection.
#cross-references MIM:95286865
#accession  I55617
#status     preliminary; translated from GB/EMBL/DBD

```

```

#molecule_type mRNA
#residues     1-147 #label RES
#cross-references EMBL:U17919; NID:g972908; CDS_PID:g972909
SUMMARY       #length 147 #molecular-weight 16827 #checksum 2715
Query Match   77.8%; Score 56; DB 14; Length 147;
Best Local Similarity 88.9%; Pred. No. 1.85e-01;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 24 ldginkhfl 32
|||||
QY 2 LDGGNKHFL 10

RESULT 5
ENTRY   A49179 #type fragment
TITLE   melanoma antigen homolog rpel - bovine (fragment)
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE    19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
ACCESSIONS A49179
REFERENCE   A49179
#authors    Kim, R.Y.; Wistow, G.J.
#journal    Exp. Eye Res. (1992) 55:657-662
#title      The cDNA RPE1 and monoclonal antibody HMB-50 define gene
              products preferentially expressed in retinal pigment
              epithelium.
#cross-references MIM:93122163
#accession  A49179
#status     preliminary
#molecule_type nucleic acid
#residues     1-491 #label KIM
#cross-references NCBI:122438; NCBI:122439
#experimental_source retinal pigment epithelium
#note        sequence extracted from NCBI backbone
SUMMARY       #length 491 #checksum 3125
Query Match   73.6%; Score 53; DB 14; Length 491;
Best Local Similarity 80.0%; Pred. No. 8.29e-01;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 67 aldggrnkrl 76
|||||
QY 1 ALDGGNKHFL 10

RESULT 6
ENTRY   C64458 #type complete
TITLE   high-affinity branched-chain amino acid transport ATP-binding
              protein - Methanococcus jannaschii
ORGANISM #formal_name Methanococcus jannaschii
DATE    13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
ACCESSIONS C64458
REFERENCE   A64300
#authors    Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
              R.D.; Sutton, G.G.; Blake, J.A.; Fitzgerald, L.M.; Clayton,
              R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
              Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
              Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glödek, A.;
              Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann,
              J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,
              J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
              K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
              H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
#journal    Science (1996) 273:1058-1073
#title      Complete genome sequence of the methanogenic archaeon,
              Methanococcus jannaschii.
#accession  C64458
#status     preliminary; nucleic acid sequence not shown;
              translation not shown
#molecule_type DNA
#residues     1-191 #label BUL

```

```

##cross-references GB:I77117; TIGR:WJ1268; CDS_PID:gl5111274
GENETICS
#map_position FOR1212818-1213393
#start_codon TTG
SUMMARY
#length 191 #molecular-weight 21466 #checksum 641

Query Match 70.8%; Score 51; DB 10; Length 191;
Best Local Similarity 50.0%; Pred. No. 2.19e+00;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 87 tlggggrqfl 96
:||||:|
Qy 1 ALDGGNKHFL 10

RESULT 7
ENTRY C64074 #type complete
TITLE fructose-bisphosphate aldolase (fba) homolog - Haemophilus
influenzae (strain Rd KW20)
ORGANISM #formal_name Haemophilus influenzae
DATE 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
03-May-1996
ACCESSIONS C64074
REFERENCE A64000
#authors Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;
Kirkness, E.F.; Kervilave, A.R.; Bult, C.J.; Tomb, J.F.;
Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.;
FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton,
M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saudek,
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann,
J.L.; Geoghagen, N.S.M.; Gnehm, C.L.; McDonald, L.A.;
Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
#journal Science (1995) 269:496-512
#title Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.
#accession C64074
#status preliminary
#molecule_type DNA
#residues 1-359 #label TIGR
##cross-references GB:I42023; TIGR:HI0524
#note named as homolog to a protein from Escherichia coli
CLASSIFICATION #superfamily fructose-bisphosphate aldolase II
SUMMARY #length 359 #molecular-weight 39339 #checksum 2903

Query Match 69.4%; Score 50; DB 6; Length 359;
Best Local Similarity 75.0%; Pred. No. 3.52e+00;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 124 ldagekhf 131
|||||
Qy 2 LDGKNKHF 9

RESULT 8
ENTRY ADEC2A #type complete
TITLE fructose-bisphosphate aldolase (EC 4.1.2.13) II - Escherichia
coli
ALTERNATE_NAMES #formal_name Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change
30-Jun-1993
ACCESSIONS S02177; A38058; S04734
REFERENCE S02177
#authors Alefounder, P.R.; Baldwin, S.A.; Perham, R.N.; Short, N.J.
#journal Biochem. J. (1989) 257:529-534
#title Cloning, sequence analysis and over-expression of the gene
for the class II fructose 1,6-bisphosphate aldolase of
Escherichia coli.
#cross-references MUID:89193446
#accession S02177
#molecule_type DNA
##residues 1-359 #label ALE
##cross-references EMBL:X14682
#accession A38058
#molecule_type protein
#residues 165-167; 213-216; 277-285; 337-345 #label ALE2
GENETICS
#gene fba; fda
#map_position 63 min
CLASSIFICATION #superfamily fructose-bisphosphate aldolase II
KEYWORDS aldehyde-lyase; carbon-carbon lyase; gluconeogenesis;
glycolysis; homodimer; pentose phosphate pathway; zinc
FEATURE
2-359 #product fructose-bisphosphate aldolase #status
experimental #label MAT
SUMMARY #length 359 #molecular-weight 39147 #checksum 6863

Query Match 69.4%; Score 50; DB 2; Length 359;
Best Local Similarity 75.0%; Pred. No. 3.52e+00;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 124 ldagekhf 131
|||||
Qy 2 LDGKNKHF 9

RESULT 9
ENTRY JS0202 #type complete
TITLE esterase - Acinetobacter calcoaceticus
ORGANISM #formal_name Acinetobacter calcoaceticus
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
23-Mar-1995
ACCESSIONS JS0202
REFERENCE JS0202
#authors Reddy, P.G.; Allon, R.; Mevarech, M.; Mendelovitz, S.; Sato,
Y.; Gutnick, D.L.
#journal Gene (1989) 76:145-152
#title Cloning and expression in Escherichia coli of an
esterase-coding gene from the oil-degrading bacterium
Acinetobacter calcoaceticus RAG-1.
#cross-references MUID:89306649
#accession JS0202
#molecule_type DNA
#residues 1-290 #label RED
##experimental_source RAG-1
#note the authors translated the codon GAT for residue 110 as
Asn and TGG for residue 256 as Tyr
COMMENT Acinetobacter calcoaceticus RAG-1 has the ability to grow on simple
triglycerides such as triacetin; mutants defective in esterase
lose this ability.
GENETICS
#gene est
SUMMARY #length 290 #molecular-weight 32711 #checksum 1173

Query Match 68.1%; Score 49; DB 8; Length 290;
Best Local Similarity 60.0%; Pred. No. 5.64e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 186 alqagikhyl 195
|||||
Qy 1 ALDGGNKHFL 10

RESULT 10
ENTRY S53472 #type complete
TITLE hypothetical protein YAR064w - yeast (Saccharomyces
cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 05-May-1995 #sequence_revision 01-Sep-1995 #text_change
12-Apr-1996
ACCESSIONS S53472
REFERENCE S53458
#authors Bussey, H.; Keng, T.; Storms, R.K.; Vo, D.; Zhong, W.;
Fortin, N.; Barton, A.B.; Kaback, D.B.; Clark, M.W.

```



```

#submission submitted to the EMBL Data Library, February 1994
#description Sequencing of chromosome I of Saccharomyces cerevisiae:
#analysis of the 52Kbp CDC15- FLO1-PHO11-YAR074 region.
#accession S53472
#molecule_type DNA
#residues 1-99 ##label BUS
##cross-references EMBL:L28920
GENETICS
#map_position 1R
SUMMARY
#length 99 #molecular-weight 11139 #checksum 3589
Query Match 66.7%; Score 48; DB 11; Length 99;
Best Local Similarity 50.0%; Pred. NO. 8.96e+00;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 66 slegsrkdf1 75
:|:|:|
QY 1 ALDGGNKHFL 10

RESULT 11
ENTRY S30821 #type complete
TITLE hypothetical protein 6 - yeast (Saccharomyces cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 28-May-1993 #sequence_revision 28-May-1993 #text_change
28-May-1993
ACCESSIONS S30821
REFERENCE S30812
#authors Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.;
#submission submitted to the EMBL Data Library, February 1993
#accession S30821
#molecule_type DNA
#residues 1-118 ##label MUL
##cross-references EMBL:L10718
GENETICS
#map_position 5
SUMMARY
#length 118 #molecular-weight 13420 #checksum 8304
Query Match 66.7%; Score 48; DB 11; Length 118;
Best Local Similarity 70.0%; Pred. NO. 8.96e+00;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 96 slpggnkgf1 105
:|:|:|
QY 1 ALDGGNKHFL 10

RESULT 12
ENTRY A57362 #type fragment
TITLE gyrb protein - Streptococcus pneumoniae (fragment)
ORGANISM #formal_name Streptococcus pneumoniae
DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
26-Jul-1996
ACCESSIONS A57362
REFERENCE A57362
#authors Munoz, R.; Bustamante, M.; de la Campa, A.G.
#journal J. Bacteriol. (1995) 177:4166-4170
#title Ser-127-to-Leu substitution in the DNA gyrase B subunit of
Streptococcus pneumoniae is implicated in novobiocin
resistance.
#cross-references MUID:95332230
#accession A57362
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-144 ##label RES
##cross-references EMBL:X83917; NID:g1052802; CDS_PID:g1052803
GENETICS
#gene gyrb
SUMMARY
#length 144 #checksum 6405
Query Match 65.3%; Score 47; DB 10; Length 144;

```

```

Best Local Similarity 55.6%; Pred. NO. 1.41e+01;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 49 lngggrhfl 57
:|:|:|
QY 2 LDGGNKHFL 10

RESULT 13
ENTRY S50964 #type complete
TITLE hypothetical protein YLL056c - yeast (Saccharomyces
cerevisiae)
ALTERNATE_NAMES hypothetical protein I0575
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 11-Feb-1995 #sequence_revision 12-May-1995 #text_change
19-Jul-1996
ACCESSIONS S50964; S64808
REFERENCE S50950
#authors Wedler, H.; Wambutt, R.
#submission submitted to the EMBL Data Library, January 1995
#description Sequence of a 37 kb DNA fragment from chromosome XII of
Saccharomyces cerevisiae including the subtelomeric region
of the left arm.
#accession S50964
#molecule_type DNA
#residues 1-298 ##label WED
##cross-references EMBL:247973
REFERENCE S64792
#authors Wedler, H.; Wedler, E.; Scharfe, M.; Wambutt, R.
#submission submitted to the Protein Sequence Database, May 1996
#accession S64808
#molecule_type DNA
#residues 1-298 ##label WEW
##cross-references EMBL:273161
##experimental_source strain S288C
GENETICS
#map_position 12L
SUMMARY
#length 298 #molecular-weight 32148 #checksum 2727
Query Match 65.3%; Score 47; DB 11; Length 298;
Best Local Similarity 60.0%; Pred. NO. 1.41e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 99 slkgsnkpfl 108
:|:|:|
QY 1 ALDGGNKHFL 10

RESULT 14
ENTRY S33473 #type complete
TITLE interleukin-1 receptor type 2 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
20-Feb-1995
ACCESSIONS S33473
REFERENCE S33473
#authors Bristulif, J.; Gatti, S.; Bartfai, T.
#submission submitted to the EMBL Data Library, May 1993
#description The rat insulinoma beta-cell line RINm5F express mRNA for the
type 2 interleukin-1 receptor.
#accession S33473
#status preliminary
#molecule_type mRNA
#residues 1-416 ##label BRI
##cross-references EMBL:222812
SUMMARY
#length 416 #molecular-weight 46353 #checksum 3686
Query Match 65.3%; Score 47; DB 14; Length 416;
Best Local Similarity 77.8%; Pred. NO. 1.41e+01;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 188 ldkgnkxfl 196
:|:|:|

```

```
QY      2 LDGKNKHL 10

RESULT  15
ENTRY   QRRBVD      #type complete
TITLE   VLDL receptor precursor - rabbit
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
         rabbit
DATE    31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
        01-Dec-1995
ACCESSIONS A46286
REFERENCE  A46286
AUTHORS   Takahashi, S.; Kawarabayashi, Y.; Nakai, T.; Sakai, J.;
          Yamamoto, T.
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. (1992) 89:9252-9256
TITLE     Rabbit very low density lipoprotein receptor: a low density
          lipoprotein receptor-like protein with distinct ligand
          specificity.
#cross-references MUID:93028442
#accession      A46286
#molecule_type mRNA
#residues      1-873 #label TAK
#cross-references NCBI:115853; NCBI:115854; DBJ:D11100
#note          sequence extracted from NCBI backbone
COMMENT      This protein is abundant in heart, muscle, and adipose tissue.
CLASSIFICATION #superfamily LDL receptor; EGF homology; LDL receptor
               ligand-binding repeat homology; LDL receptor
               YWTD-containing repeat homology; LDL receptor/EGF precursor
               homology
KEYWORDS      cholesterol; coated pits; duplication; endocytosis;
               glycoprotein; lipid transport; receptor; transmembrane
               protein; VLDL

FEATURE
1-27      #domain signal sequence #status predicted #label SIG
28-873    #product VLDL receptor #status predicted #label MAT\
28-797    #domain extracellular #status predicted #label EXT\
33-67     #domain LDL receptor ligand-binding repeat homology
          #label LDL9\
72-108    #domain LDL receptor ligand-binding repeat homology
          #label LDL1\
113-149   #domain LDL receptor ligand-binding repeat homology
          #label LDL2\
154-188   #domain LDL receptor ligand-binding repeat homology
          #label LDL3\
193-229   #domain LDL receptor ligand-binding repeat homology
          #label LDL4\
239-273   #domain LDL receptor ligand-binding repeat homology
          #label LDL5\
278-312   #domain LDL receptor ligand-binding repeat homology
          #label LDL6\
318-355   #domain LDL receptor ligand-binding repeat homology
          #label LDL7\
360-749   #domain LDL receptor/EGF precursor homology #label LDEG\
360-394   #domain EGF homology #label EG1\
400-434   #domain EGF homology #label EG2\
441-480   #domain LDL receptor YWTD-containing repeat homology
          #label YW1\
481-524   #domain LDL receptor YWTD-containing repeat homology
          #label YW2\
525-567   #domain LDL receptor YWTD-containing repeat homology
          #label YW3\
568-611   #domain LDL receptor YWTD-containing repeat homology
          #label YW4\
612-654   #domain LDL receptor YWTD-containing repeat homology
          #label YW5\
655-697   #domain LDL receptor YWTD-containing repeat homology
          #label YW6\
706-749   #domain EGF homology #label EG3\
751-790   #region clustered O-linked oligosaccharides\
798-819   #domain transmembrane #status predicted #label TMM\
820-873   #domain intracellular #status predicted #label INT\
834-838   #region coated-pit mediated internalization signal\
151,765,781 #binding_site carbohydrate (Asn) (covalent) #status
```

```

360-371,367-380,
382-394,400-410,
406-419,421-434,
706-719,715-734,
736-749
SUMMARY      #length 873 #molecular_weight 96280 #checksum 5713
              #disulfide_bonds #status predicted

Query Match 65.3%; Score 47; DB 2; Length 873;
Best Local Similarity 50.0%; Pred. No. 1.41e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 541 tldgkkrkfl 550
QY 1 ALDGGNKHL 10
:||||: ||
:||||: ||

Search completed: Tue Jun 10 11:24:32 1997
Job time : 10 secs.

predicted\
```

\*\*\*\*\*

WQSEELH (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:23:55 1997; MasPar time 2.07 Seconds  
Tabular output not generated. 102.659 Million cell updates/sec

Title: >US-08-231-565A-33  
Description: (1-10) from US08231565A.pep  
Perfect score: 72  
Sequence: 1 ALDGNKHF 10

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 23.690; Variance 23.676; scale 1.001

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	72	100.0	668	PM17_HUMAN	PMEL 17 PROTEIN PRECU	6.48e-07
2	56	77.8	147	A1FL_RAT	ALLOGRAFT INFLAMMATOR	2.20e-02
3	52	72.2	455	P2X5_RAT	P2X PURINOCEPTOR 5 (A	2.35e-01
4	50	69.4	358	1 ALF_ECOLI	FRUCTOSE-BISPHOSPHATE	7.35e-01
5	50	69.4	359	1 ALF_HAIEIN	FRUCTOSE-BISPHOSPHATE	7.35e-01
6	49	68.1	303	EST_ACICA	ESTERASE (EC 3.1.1.-)	1.28e+00
7	48	66.7	99	1 YAN4_YEAST	HYPOTHETICAL 11.1 KD	2.22e+00
8	48	66.7	410	5 ILIS_MOUSE	INTERLEUKIN-1 RECEPTO	2.22e+00
9	47	65.3	416	5 ILIS_MOUSE	INTERLEUKIN-1 RECEPTO	3.81e+00
10	47	65.3	873	5 LDVR_RABIT	VERY LOW-DENSITY LIPO	3.81e+00
11	47	65.3	873	5 LDVR_MOUSE	VERY LOW-DENSITY LIPO	3.81e+00
12	46	63.9	179	5 HTRC_ECOLI	HEAT SHOCK PROTEIN C.	6.47e+00
13	46	63.9	401	11 Y148_METSM	INSERTION ELEMENT ISM	6.47e+00
14	46	63.9	489	1 6PGD_YEAST	6-PHOSPHOGUCCONATE DE	6.47e+00
15	46	63.9	492	1 6PGD_YEAST	PROBABLE 6-PHOSPHOGLU	6.47e+00
16	46	63.9	1099	3 CYA7_MOUSE	ADENYLATE CYCLASE TY	6.47e+00
17	45	62.5	217	11 YG18_HAIEIN	HYPOTHETICAL ABC TRAN	1.09e+01
18	45	62.5	245	1 BIOW_BACSH	6-CARBOXYHEXANONATE--C	1.09e+01
19	45	62.5	316	11 Y0A4_CAEEL	HYPOTHETICAL 36.5 KD	1.09e+01
20	45	62.5	354	1 BGAT_HUMAN	FUCOSYLGLYCOPROTEIN A	1.09e+01
21	45	62.5	368	4 GATR_BOVIN	N-ACETYLACTOSAMINIDE	1.09e+01
22	45	62.5	478	1 BH5_MOUSE	BRAIN PROTEIN H5.	1.09e+01

23	45	62.5	516	4 GLC5_SOYBN	GLYCININ PRECURSOR (C	1.09e+01
24	45	62.5	562	4 GLC4_SOYBN	GLYCININ G4 PRECURSOR	1.09e+01
25	45	62.5	873	5 LDVR_HUMAN	VERY LOW-DENSITY LIPO	1.09e+01
26	45	62.5	873	5 LDVR_RAT	VERY LOW-DENSITY LIPO	1.09e+01
27	45	62.5	1106	1 ACVY_CAEEL	PROBABLE ATP-CITRATE	1.09e+01
28	44	61.1	128	10 VG64_BPM15	GENE 64 PROTEIN (GP64	1.81e+01
29	44	61.1	151	11 YF12_F0WPV	HYPOTHETICAL PROTEIN	1.81e+01
30	44	61.1	210	11 YAT1_YEAST	HYPOTHETICAL 24.4 KD	1.81e+01
31	44	61.1	296	9 THTR_HUMAN	THIOSULFATE SULFURTRA	1.81e+01
32	44	61.1	393	6 METK_LYCES	S-ADENOSYLMETHIONINE	1.81e+01
33	44	61.1	449	10 Y425_MYCGE	PROBABLE RNA HELICASE	1.81e+01
34	44	61.1	467	1 6PGD_BACLI	6-PHOSPHOGUCCONATE DE	1.81e+01
35	44	61.1	488	1 6PGD_BACSU	6-PHOSPHOGUCCONATE DE	1.81e+01
36	44	61.1	623	8 RSD1_YEAST	RECESSIVE SUPPRESSOR	1.81e+01
37	44	61.1	896	1 APCE_SYNY4	PHYCOBILISOME 100.5 K	1.81e+01
38	44	61.1	995	11 YPDI_CAEEL	HYPOTHETICAL 111.2 KD	1.81e+01
39	44	61.1	1086	6 MYST_HUMAN	MYOSIN HEAVY CHAIN, S	1.81e+01
40	44	61.1	1474	1 A2MG_HUMAN	ALPHA-2-MACROGLOBULIN	1.81e+01
41	44	61.1	1894	3 FAS2_YEAST	FATTY ACID SYNTHASE,	1.81e+01
42	43	59.7	553	10 VGLF_NDVH4	FUSION GLYCOPROTEIN P	2.99e+01
43	43	59.7	553	10 VGLF_NDVM	FUSION GLYCOPROTEIN P	2.99e+01
44	43	59.7	553	10 VGLF_NDVI	FUSION GLYCOPROTEIN P	2.99e+01
45	43	59.7	553	10 VGLF_NDVT	FUSION GLYCOPROTEIN P	2.99e+01

ALIGNMENTS

RESULT	1	PM17_HUMAN	STANDARD;	PRT;	668 AA.
AC	P40967;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE: 92021023.				
RA	KWON B.S., CHINTAMANANI C., KOZAK C.A., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,				
RA	KIM K.-K.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).				
CC	-!- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	-!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; G190106; -.				
DR	MIM; 155550; -.				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23	POTENTIAL.	
FT	CHAIN	24	668	PMEL 17 PROTEIN.	
FT	TRANSMEM	575	595	POTENTIAL.	
FT	TRANSMEM	603	623	POTENTIAL.	
FT	DOMAIN	217	307	PKD.	
FT	DOMAIN	315	444	10 X 13 AA TANDEM REPEATS.	
FT	REPEAT	315	327	1.	
FT	REPEAT	328	340	2.	
FT	REPEAT	341	353	3.	
FT	REPEAT	354	366	4.	
FT	REPEAT	367	379	5.	
FT	REPEAT	380	392	6.	
FT	REPEAT	393	405	7.	
FT	REPEAT	406	418	8.	
FT	REPEAT	419	431	9.	
FT	REPEAT	432	444	10.	
FT	CARBOHYD	81	81	POTENTIAL.	
FT	CARBOHYD	106	106	POTENTIAL.	
FT	CARBOHYD	111	111	POTENTIAL.	
FT	CARBOHYD	321	321	POTENTIAL.	
FT	CARBOHYD	568	568	POTENTIAL.	
FT	SEQUENCE	668 AA;	70992 MW;	6E8E1AFO CRC32;	

Query Match 100.0%; Score 72; DB 7; Length 668;  
 Best Local Similarity 100.0%; Pred. No. 6.48e-07;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 224 aldggnkhfl 233  
 |||||  
 QY 1 LDGNGKHFL 10

RESULT 2 STANDARD; PRT; 147 AA.  
 ID AIF1\_RAT  
 AC P5009;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE ALLOGRAFT INFLAMMATORY FACTOR-1 (AIF-1).  
 GN AIF1.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LEWIS; TISSUE-CARDIAC;  
 RX MEDLINE; 95286865.  
 RA UTANS U., ARCECI R.J., YAMASHITA Y., RUSSELL M.E.;  
 RL J. CLIN. INVEST. 95:2954-2962(1995).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN MACROPHAGE ACTIVATION AND FUNCTION.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY AND PERSISTENTLY IN  
 CC CHRONICALLY REJECTING CARDIAC ALLOGRAFTS BUT IS ABSENT IN CARDIAC  
 CC SYNGRAFTS AND HOST HEARTS.  
 CC -1- SIMILARITY: SOME, TO EF-HAND CALCIUM BINDING PROTEINS.  
 CC -1- SIMILARITY: STRONG, TO BALLOON ANGIOPLASTY RESPONSIVE TRANSCRIPT 1  
 CC (BAT-1).  
 DR EMBL; U17919; G972909; -.  
 KW CALCIUM-BINDING.  
 FT CA\_BIND 58 69 POTENTIAL.  
 FT DOMAIN 94 105 ANCESTRAL CALCIUM SITE.  
 SQ SEQUENCE 147 AA; 16827 MW; A9388B0E CRC32;

Query Match 77.8%; Score 56; DB 1; Length 147;  
 Best Local Similarity 88.9%; Pred. No. 2.20e-02;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 24 ldginkhfl 32  
 ||| |||||  
 QY 2 LDGNGKHFL 10

RESULT 3 STANDARD; PRT; 455 AA.  
 ID P2X5\_RAT  
 AC P51578;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE P2X PURINORECEPTOR 5 (ATP RECEPTOR) (P2X5) (PURINERGIC RECEPTOR).  
 GN P2RX5.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COELIAC GANGLION;  
 RX MEDLINE; 96256686.  
 RA COLLO G., KAWASHIMA E., PICH E., NEIDHART S., NORTH R.A.,  
 RA SURPRENANT A., BUELL G.N.;  
 RL J. NEUROSCI. 16:2495-2507(1996).  
 CC -1- FUNCTION: RECEPTOR FOR ATP THAT ACTS AS A LIGAND GATED ION  
 CC CHANNEL.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SUBUNIT: HOMO- OR HETEROPOLYMERS (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE P2X RECEPTOR FAMILY.  
 DR EMBL; X92069; E205287; -.

KW IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; RECEPTOR; GLYCOPROTEIN.  
 FT DOMAIN 1 30  
 FT TRANSEM 31 51  
 FT DOMAIN 52 341  
 FT TRANSEM 342 362  
 FT DOMAIN 363 455  
 FT CARBOHYD 77 77  
 FT CARBOHYD 157 157  
 FT CARBOHYD 202 202  
 SQ SEQUENCE 455 AA; 51479 MW; 7EDE74C3 CRC32;

Query Match 72.2%; Score 52; DB 7; Length 455;  
 Best Local Similarity 66.7%; Pred. No. 2.35e-01;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 209 letdnkhfl 217  
 |: |||||  
 QY 2 LDGNGKHFL 10

RESULT 4 STANDARD; PRT; 358 AA.  
 ID ALF\_ECOLI  
 AC P11604;  
 DT 01-OCT-1989 (REL. 12, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE FRUCTOSE-BISPHOSPHATE ALDOLASE (EC 4.1.2.13).  
 GN FBA OR FDA.  
 OS ESCHERICHIA COLI.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / CS520;  
 RX MEDLINE; 89313302.  
 RA ALEFINDER P.R., BALDWIN S.A., PERHAM S.A., SHORT N.J.;  
 RL MOL. MICROBIOL. 3:723-732(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-26.  
 RX MEDLINE; 89193446.  
 RA ALEFINDER P.R., BALDWIN S.A., PERHAM R.N., SHORT N.J.;  
 RL BIOCHEM. J. 257:529-534(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RA PLUNKETT G. III;  
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [4]  
 RP SEQUENCE OF 1-12.  
 RC STRAIN-K12 / EMG2;  
 RA LINK A.J.;  
 RL SUBMITTED (OCT-1994) TO THE SWISS-PROT DATA BANK.  
 RN [5]  
 RP ZINC-LIGANDS, AND MUTAGENESIS.  
 RX MEDLINE; 93170474.  
 RA BERRY A., MARSHALL K.E.;  
 RL FEBS LETT. 318:11-16(1993).  
 CC -1- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE - GLYCERONE-  
 CC PHOSPHATE + D-GLYCERALDEHYDE 3-PHOSPHATE.  
 CC -1- PATHWAY: SIXTH STEP IN GLYCOLYSIS.  
 CC -1- COFACTOR: ZINC.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS II FRUCTOSE-BISPHOSPHATE ALDOLASES.  
 DR EMBL; X14436; G41423; -.  
 DR EMBL; U28377; G882454; -.  
 DR PIR; S02177; ADEC2A.  
 DR ECOGENE; EG10282; FBA.  
 DR PROSITE; PS00602; ALDOLASE\_CLASS\_II\_1.  
 DR PROSITE; PS00806; ALDOLASE\_CLASS\_II\_2.  
 KW LYASE; GLYCOLYSIS; ZINC.  
 FT INIT\_MET 0 0  
 FT METAL 107 107 ZINC.  
 FT METAL 110 110 ZINC.

```

FT MUTAGEN 107 107 H->A: LOSS OF ACTIVITY.
FT MUTAGEN 110 110 H->A: LOSS OF ACTIVITY.
FT MUTAGEN 111 111 C->A: PARTIAL LOSS OF ACTIVITY.
SQ SEQUENCE 358 AA; 39016 MW; 2876D61F CRC32;

Query Match 69.4%; Score 50; DB 1; Length 358;
Best Local Similarity 75.0%; Pred. No. 7.35e-01; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 0;

Db 123 ldagekhf 130
||:|:|:|
QY 2 LDGKNKF 9

RESULT 5
ID ALF HAEIN STANDARD; PRT; 359 AA.
AC P44429;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE FRUCTOSE-BISPHOSPHATE ALDOLASE (EC 4.1.2.13).
GN FBA OR HI0524.
OS HAEMOPHILUS INFLUENZAE.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
CC PASTEURILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB G.-F., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RL SCIENCE 269:496-512(1995).
CC -!- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE = GLYCERONE-
CC PHOSPHATE + D-GLYCERALDEHYDE 3-PHOSPHATE.
CC -!- PATHWAY: SIXTH STEP IN GLYCOLYSIS.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO CLASS II FRUCTOSE-BISPHOSPHATE ALDOLASES.
DR EMBL; L45165; G1003932; -.
DR EMBL; U32734; G925549; -.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2.
KW LYASE; GLYCOLYSIS; ZINC.
FT METAL 108 108 ZINC (BY SIMILARITY).
FT METAL 111 111 ZINC (BY SIMILARITY).
SQ SEQUENCE 359 AA; 39339 MW; B7FC8E1A CRC32;

Query Match 69.4%; Score 50; DB 1; Length 359;
Best Local Similarity 75.0%; Pred. No. 7.35e-01; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 0;

Db 124 ldagekhf 131
||:|:|:|
QY 2 LDGKNKF 9

RESULT 6
ID EST ACICA STANDARD; PRT; 303 AA.
AC P18773;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE ESTERASE (EC 3.1.1.-).
GN EST.
OS ACINETOBACTER CALCOACETICUS.

```

```

OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC NEISSERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RAG-1;
RX MEDLINE; 89306649.
RA REDDY P.G., ALLON R., MEVARECH M., MENDELOVITZ S., SATO Y.,
RA GUTNICK D.L.;
RL GENE 76:145-152(1989).
RN [2]
RP REVISIONS.
RA GUTNICK D.L.;
RL SUBMITTED (JUL-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- ACINETOBACTER CALCOACETICUS RAG-1 HAS THE ABILITY TO GROW ON
CC SIMPLE TRIGLYCERIDES SUCH AS TRIACETIN; MUTANTS DEFECTIVE IN
CC ESTERASE LOSE THIS ABILITY, BUT RETAIN THE ABILITY TO GROW ON
CC ACETATE.
CC -!- SUBCELLULAR LOCATION: EXTERNAL TO THE CYTOPLASMIC MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE "GDHG" FAMILY OF LIPOLYTIC ENZYMES.
DR EMBL; M24890; G303953; -.
DR PIR; JS0202; JS0202.
KW HYDROLASE; SERINE ESTERASE.
FT ACT_SITE 79 79 POTENTIAL.
FT ACT_SITE 149 149 POTENTIAL.
SQ SEQUENCE 303 AA; 33911 MW; 249D989E CRC32;

Query Match 68.1%; Score 49; DB 3; Length 303;
Best Local Similarity 60.0%; Pred. No. 1.28e+00;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 203 alqagikhyl 212
||:|:|:|
QY 1 ALDGGNKF 10

RESULT 7
ID YAN4 YEAST STANDARD; PRT; 99 AA.
AC P39563;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 11.1 KD PROTEIN IN FLO1-PHO11 INTERGENIC REGION.
GN YAR064W.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA BUSSEY H., KENG T., STORMS R.K., VO D., ZHONG W., FORTIN N.,
RA BARTON A.B., KABACK D.B., CLARK M.W.;
RL SUBMITTED (FEB-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RX MEDLINE; 95249563.
RA BUSSEY H., KABACK D.B., ZHONG W., VO D.T., CLARK M.W., FORTIN N.,
RA HALL J., OUELLETTE B.F., KENG T., BARTON A.B., SU Y., DAVIES C.K.,
RA STORMS R.K.;
RL PROC. NATL. ACAD. SCI. U.S.A. 92:3809-3813(1995).
DR EMBL; L28920; G456151; -.
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT TRANSMEM 74 90 POTENTIAL.
SQ SEQUENCE 99 AA; 11139 MW; 477D420E CRC32;

Query Match 66.7%; Score 48; DB 11; Length 99;
Best Local Similarity 50.0%; Pred. No. 2.22e+00;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 66 slegrskdfl 75
||:|:|:|
QY 1 ALDGGNKF 10

```

```

RESULT 8
ID IL1S_MOUSE STANDARD; PRT; 410 AA.
AC P27931;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE INTERLEUKIN-1 RECEPTOR, TYPE II PRECURSOR (IL-1R-2).
GN IL1RB OR IL-1R2.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92007725.
RA MCMAHAN C.J., SLACK J.L., MOSLEY B., COSMAN D., LUPTON S.D.,
RA BRUNTON L.L., GRUBIN C.E., WIGNALL J.M., JENKINS N.A., BRANNAN C.I.,
RA COPELAND N.G., HUEBNER K., CROCE C.M., CANNIZZARRO L.A., BENJAMIN D.,
RA DOWER S.K., SPRIGGS M.K., SIMS J.E.;
RL EMBO J. 10:2821-2832(1991).
CC -1- FUNCTION: RECEPTOR FOR INTERLEUKIN-1 ALPHA (IL-1A), BETA (IL-1B),
CC AND INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN (IL-1RA).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC THREE C2-LIKE DOMAINS.
DR EMBL; X59769; G52671; -.
KW IMMUNOGLOBULIN FOLD; RECEPTOR; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.
FT SIGNAL 1 13
FT CHAIN 14 410
FT DOMAIN 14 355
FT TRANSMEM 356 381
FT DOMAIN 382 410
FT DOMAIN 57 127
FT DOMAIN 157 226
FT DOMAIN 263 345
FT DISULFID 64 120
FT DISULFID 164 219
FT DISULFID 270 338
FT CARBOHYD 124 124
FT CARBOHYD 208 208
FT CARBOHYD 231 231
FT CARBOHYD 289 289
SQ SEQUENCE 410 AA; 45645 MW; 54F5A526 CRC32;

Query Match 66.7%; Score 48; DB 5; Length 410;
Best Local Similarity 77.8%; Pred. No. 2.22e+00;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 188 ldkgnkef1 196
||| ||| |||
QY 2 LDGNGKHFL 10

RESULT 9
ID IL1S_RAT STANDARD; PRT; 416 AA.
AC P43303;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE INTERLEUKIN-1 RECEPTOR, TYPE II PRECURSOR (IL-1R-2).
GN IL1RB OR IL-1R2.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95035882.
RA BRISTOLF J., GATTI S., MALINOWSKY D., BJORK L., SUNDGREN A.K.,
RA BARTAI T.;
RL EUR. CYTOKINE NETW. 5:319-330(1994).
CC -1- FUNCTION: RECEPTOR FOR INTERLEUKIN-1 ALPHA (IL-1A), BETA (IL-1B),
CC AND INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN (IL-1RA).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS

```

```

CC THREE C2-LIKE DOMAINS.
DR EMBL; Z22812; G311408; -.
KW IMMUNOGLOBULIN FOLD; RECEPTOR; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.
FT SIGNAL 1 13
FT CHAIN 14 416
FT DOMAIN 14 355
FT TRANSMEM 356 381
FT DOMAIN 382 416
FT DOMAIN 57 127
FT DOMAIN 157 226
FT DOMAIN 263 345
FT DISULFID 64 120
FT DISULFID 164 219
FT DISULFID 270 338
FT CARBOHYD 124 124
FT CARBOHYD 208 208
FT CARBOHYD 231 231
FT CARBOHYD 289 289
SQ SEQUENCE 416 AA; 46353 MW; 378EFB34 CRC32;

Query Match 65.3%; Score 47; DB 5; Length 416;
Best Local Similarity 77.8%; Pred. No. 3.81e+00;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 188 ldkgnkef1 196
||| ||| |||
QY 2 LDGNGKHFL 10

RESULT 10
ID LDVR_RABIT STANDARD; PRT; 873 AA.
AC P35953;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE VERY LOW-DENSITY LIPOPROTEIN RECEPTOR PRECURSOR (VLDL RECEPTOR).
GN VLDLR.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; LAGOMORPHA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93028442.
RA TAKARASHI S., KAWARABAYASI Y., NAKAI T., SAKAI J., YAMAMOTO T.;
RL PROC. NATL. ACAD. SCI. U.S.A. 89:9252-9256(1992).
CC -1- FUNCTION: BINDS VLDL AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS.
CC IN ORDER TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST
CC FIRST CLUSTER INTO CLATHRIN-COATED PITS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ABUNDANT IN HEART, MUSCLE, AND ADIPOSE
CC TISSUE.
CC -1- SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
DR EMBL; D11100; -. NOT_ANNOTATED_CDS.
DR PIR; A46286; QRRBVD.
KW GLYCOPROTEIN; VLDL; CHOLESTEROL METABOLISM; LIPID TRANSPORT;
KW ENDOCYTOSIS; COATED PITS; TRANSMEMBRANE; RECEPTOR; SIGNAL;
KW EGF-LIKE DOMAIN; REPEAT.
FT SIGNAL 1 27
FT CHAIN 28 873
FT DOMAIN 28 797
FT TRANSMEM 798 819
FT DOMAIN 820 873
FT DOMAIN 31 69
FT DOMAIN 70 110
FT DOMAIN 111 151
FT DOMAIN 152 190
FT DOMAIN 191 231
FT DOMAIN 237 275
FT DOMAIN 276 314
FT DOMAIN 316 355
FT DOMAIN 356 395

```

```

FT DOMAIN 396 435 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 439 480 LDL-RECEPTOR CLASS B 1.
FT REPEAT 481 524 LDL-RECEPTOR CLASS B 2.
FT REPEAT 525 567 LDL-RECEPTOR CLASS B 3.
FT REPEAT 568 611 LDL-RECEPTOR CLASS B 4.
FT REPEAT 612 654 LDL-RECEPTOR CLASS B 5.
FT REPEAT 655 696 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 702 750 EGF-LIKE 3.
FT DOMAIN 751 790 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT SITE 837 877 CRITICAL FOR ENDOCYTOSIS.
FT DISULFID 33 45 BY SIMILARITY.
FT DISULFID 40 58 BY SIMILARITY.
FT DISULFID 52 67 BY SIMILARITY.
FT DISULFID 72 84 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 91 108 BY SIMILARITY.
FT DISULFID 113 127 BY SIMILARITY.
FT DISULFID 120 140 BY SIMILARITY.
FT DISULFID 134 149 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 161 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 200 218 BY SIMILARITY.
FT DISULFID 212 229 BY SIMILARITY.
FT DISULFID 239 251 BY SIMILARITY.
FT DISULFID 246 264 BY SIMILARITY.
FT DISULFID 258 273 BY SIMILARITY.
FT DISULFID 278 290 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 297 312 BY SIMILARITY.
FT DISULFID 318 331 BY SIMILARITY.
FT DISULFID 326 344 BY SIMILARITY.
FT DISULFID 338 355 BY SIMILARITY.
FT DISULFID 360 371 BY SIMILARITY.
FT DISULFID 367 380 BY SIMILARITY.
FT DISULFID 382 394 BY SIMILARITY.
FT DISULFID 400 410 BY SIMILARITY.
FT DISULFID 406 419 BY SIMILARITY.
FT DISULFID 421 434 BY SIMILARITY.
FT DISULFID 434 451 BY SIMILARITY.
FT DISULFID 451 465 BY SIMILARITY.
FT DISULFID 465 479 BY SIMILARITY.
FT CARBOHYD 151 151 POTENTIAL.
FT CARBOHYD 765 765 POTENTIAL.
FT CARBOHYD 781 781 POTENTIAL.
SQ SEQUENCE 873 AA; 96280 MW; 90C470C5 CRC32;

Query Match 65.3%; Score 47; DB 5; Length 873;
Best Local Similarity 50.0%; Pred. No. 3.81e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 541 tldgarkrkl 550
QY 1 ALDGGNKHKL 10

RESULT 11
ID LDVR_MOUSE STANDARD; PRT; 873 AA.
AC P98156;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE VERY LOW-DENSITY LIPOPROTEIN RECEPTOR PRECURSOR (VLDL RECEPTOR).
GN VLDLR.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=HEART;
RX MEDLINE; 95010090.
RA OKA K., ISHIMURA-OKA K., CHU M.J., SULLIVAN M., KRUSHKAL J.,

```

```

RA LI W.H., CHAN L.;
RL EUR. J. BIOCHEM. 224:975-982(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SKETAL MUSCLE;
RX MEDLINE; 94283285.
RA GAFVELS M.E., PAAVOLA L.G., BOYD C.O., NOLAN P.M., WITTMACK F.,
RX CHAWLA A., LAZAR M.A., BUCAN M., ANGELIN B.O., STRAUSS J.F.;
RA ENDOCRINOLOGY 135:387-394(1994).
CC -1- FUNCTION: BINDS VLDL AND TRANSPORTS IT INTO CELLS BY ENDOCYTOSIS.
CC IN ORDER TO BE INTERNALIZED, THE RECEPTOR-LIGAND COMPLEXES MUST
CC FIRST CLUSTER INTO CLATHRIN-COATED PITS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ABUNDANT IN HEART AND MUSCLE; LESS IN KIDNEY,
CC BRAIN, OVARY, TESTIS, LUNG, ADIPOSE TISSUE.
CC -1- SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
DR EMBL; L33417; G609533; -
DR EMBL; U06670; G619647; -
KW GLYCOPROTEIN; VLDL; CHOLESTEROL METABOLISM; LIPID TRANSPORT;
KW ENDOCYTOSIS; COATED PITS; TRANSMEMBRANE; RECEPTOR; SIGNAL;
KW EGF-LIKE DOMAIN; REPEAT.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 873 VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.
FT DOMAIN 28 797 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 798 819 POTENTIAL.
FT DOMAIN 820 873 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 69 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 70 110 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 111 151 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 152 190 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 191 231 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 237 275 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 276 314 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 315 355 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 356 391 EGF-LIKE 1.
FT DOMAIN 396 431 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 439 480 LDL-RECEPTOR CLASS B 1.
FT REPEAT 481 524 LDL-RECEPTOR CLASS B 2.
FT REPEAT 525 567 LDL-RECEPTOR CLASS B 3.
FT REPEAT 568 611 LDL-RECEPTOR CLASS B 4.
FT REPEAT 612 654 LDL-RECEPTOR CLASS B 5.
FT REPEAT 655 696 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 702 750 EGF-LIKE 3.
FT DOMAIN 751 790 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT SITE 837 877 CRITICAL FOR ENDOCYTOSIS.
FT DISULFID 33 45 BY SIMILARITY.
FT DISULFID 40 58 BY SIMILARITY.
FT DISULFID 52 67 BY SIMILARITY.
FT DISULFID 72 84 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 91 108 BY SIMILARITY.
FT DISULFID 113 127 BY SIMILARITY.
FT DISULFID 120 140 BY SIMILARITY.
FT DISULFID 134 149 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 161 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 200 218 BY SIMILARITY.
FT DISULFID 212 229 BY SIMILARITY.
FT DISULFID 239 251 BY SIMILARITY.
FT DISULFID 246 264 BY SIMILARITY.
FT DISULFID 258 273 BY SIMILARITY.
FT DISULFID 278 290 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 297 312 BY SIMILARITY.
FT DISULFID 318 331 BY SIMILARITY.
FT DISULFID 326 344 BY SIMILARITY.
FT DISULFID 338 355 BY SIMILARITY.
FT DISULFID 360 371 BY SIMILARITY.
FT DISULFID 367 380 BY SIMILARITY.

```

```

FT DISULFID 382 394 BY SIMILARITY.
FT DISULFID 400 410 BY SIMILARITY.
FT DISULFID 406 419 BY SIMILARITY.
FT DISULFID 421 434 BY SIMILARITY.
FT DISULFID 706 719 BY SIMILARITY.
FT DISULFID 715 734 BY SIMILARITY.
FT DISULFID 736 749 BY SIMILARITY.
FT CARBOHYD 151 151 POTENTIAL.
FT CARBOHYD 765 765 POTENTIAL.
FT CARBOHYD 781 781 POTENTIAL.
FT CONFLICT 161 161 C -> G (IN REF. 1).
FT CONFLICT 297 297 C -> S (IN REF. 2).
SQ SEQUENCE 873 AA; 96372 MW; 80DECED7 CRC32;

Query Match 65.3%; Score 47; DB 5; Length 873;
Best Local Similarity 50.0%; Pred. No. 3.81e+00;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 541 tldgskrfkl 550
Qy 1 ALDGGNKHFL 10

RESULT 12
ID HTRC_ECOLI STANDARD; PRT; 179 AA.
AC P27375;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE HEAT SHOCK PROTEIN C.
GN HTRC.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RA OVCHINNIKOV Y.A., MONASTYRSKAYA G.S., GUBANOV V.V., GUR'YEV S.O.,
RA SALOMATINA I.S., SHUVAEVA T.M., LIPKIN V.M., SVERDLOV E.D.;
RL DOKL. BIOCHEM. 261:385-390(1981).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 94089392.
RA BLATTNER F.R., BURLAND V.D., PLUNKETT G. III, SOFIA H.J.,
RA DANIELS D.L.;
RL NUCLEIC ACIDS RES. 21:5408-5417(1993).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE; 90264341.
RA RAINA S., GEORGIOPOULOS C.;
RL J. BACTERIOL. 172:3417-3426(1990).
CC -I- INACTIVATION OF THE HTRC GENE RESULTS IN THE INABILITY TO FORM
CC COLONIES AT 42 DEGREES CELSIUS. AT 43 DEGREES CELSIUS, HTRC MUTANT
CC BACTERIA GRADUALLY LYSE, WHEREAS AT INTERMEDIATE TEMPERATURES THEY
CC FILAMENT EXTENSIVELY.
DR EMBL; V00339; -: NOT_ANNOTATED_CDS.
DR EMBL; U00006; G396328; -.
DR ECOGENE; EG11429; HTRC.
KW HEAT SHOCK.
SQ SEQUENCE 179 AA; 21131 MW; A8DCC219 CRC32;

Query Match 63.9%; Score 46; DB 5; Length 179;
Best Local Similarity 75.0%; Pred. No. 6.47e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 119 lddgdkaf 126
Qy 2 LDGGNKHFL 9

RESULT 13
ID Y148_METSM STANDARD; PRT; 401 AA.
AC P22344;

```

---

```

DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-AUG-1991 (REL. 19, LAST ANNOTATION UPDATE)
DE INSERTION ELEMENT ISM1 HYPOTHETICAL 48.3 KD PROTEIN (ORFIS).
OS METHANOBREVIBACTER SMITHII.
OC ARCHAEBACTERIA; EURYARCHAEOTA; METHANOBACTERIALES;
OC METHANOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85295498.
RA HAMILTON P.T., REEVE J.N.;
RL MOL. GEN. GENET. 200:47-59(1985).
CC -I- FUNCTION: THIS POLYPEPTIDE IS INVOLVED IN TRANSPORTATION, AND
CC SHOULD THEREFORE BIND TO NUCLEIC ACIDS.
CC -I- THE ORFIS OPEN READING FRAME REPRESENTS 87% OF THE SEQUENCE OF
CC ISM1, A MOBILE INSERTION ELEMENT.
DR EMBL; X02587; G44520; -.
DR PIR; S28653; S28653.
KW TRANSPOSABLE ELEMENT; HYPOTHETICAL PROTEIN.
SQ SEQUENCE 401 AA; 48313 MW; 13E42EEA CRC32;

Query Match 63.9%; Score 46; DB 11; Length 401;
Best Local Similarity 60.0%; Pred. No. 6.47e+00;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 114 pligenkqfl 123
Qy 1 ALDGGNKHFL 10

RESULT 14
ID 6PGD_YEAST STANDARD; PRT; 489 AA.
AC P38720;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE 6-PHOSPHOGLUCONATE DEHYDROGENASE, DECARBOXYLATING (EC 1.1.1.44).
GN GND1 OR YHR183W.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RA DESOUZA M., LOBO Z., MAITRA P.K.;
RL SUBMITTED (NOV-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RX MEDLINE; 94378003.
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
RA DU Z., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,
RA KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
RA LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
RA NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,
RA VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
RA VAUDIN M.;
RL SCIENCE 265:2077-2082(1994).
RN [3]
RP SEQUENCE OF 41-47 AND 120-131.
RC STRAIN-S288C;
RX MEDLINE; 95203288.
RA GARRELS J.I., FUTCHER B., KOBAYASHI R., LATTER G.I., SCHWENDER B.,
RA VOLPE T., WARNER J.R., MCLAUGHLIN C.S.;
RL ELECTROPHORESIS 15:1466-1486(1994).
CC -I- CATALYTIC ACTIVITY: 6-PHOSPHO-D-GLUCONATE + NADP(+) = D-RIBULOSE
CC 5-PHOSPHATE + CO(2) + NADPH.
CC -I- PATHWAY: HEXOSE MONOPHOSPHATE SHUNT.
CC -I- SIMILARITY: TO OTHER PROKARYOTIC AND EUKARYOTIC 6-PHOSPHOGLUCONATE
CC DEHYDROGENASES.
DR EMBL; Z46631; G575391; -.
DR EMBL; U00028; G458910; -.
DR EMBL; U01155; G577839; -.
DR PIR; S46671; S46671.
DR YEPD; 5460; -.

```



DR SGD; L0003125; GND1.  
DR PROSITE; PS00461; 6PGD.  
KW OXIDOREDUCTASE; PENTOSE SHUNT; NADP.  
SQ SEQUENCE 489 AA; 53543 MW; 7D30CCD2 CRC32;

Query Match 63.9%; Score 46; DB 1; Length 489;  
Best Local Similarity 75.0%; Pred. No. 6.47e+00;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 98 idggnsfhf 105  
:|||||  
QY 2 LDGKNKF 9

RESULT 15  
ID 6PG2.YEAST STANDARD; PRT; 492 AA.  
AC P5319;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE PROBABLE 6-PHOSPHOGLUCONATE DEHYDROGENASE, DECARBOXYLATING  
DE (EC 1.1.1.44).  
GN YGR256W.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA AGOSTONI CARBONE M.L., PANZERI L., MELCHIORETTO P., CARIGNANI G.,  
RA FERROLI F., FRONTALI L., MAZZONI C., RINALDI T., RUZZI M.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -1- CATALYTIC ACTIVITY: 6-PHOSPHO-D-GLUCONATE + NADP(+) = D-RIBULOSE  
CC 5-PHOSPHATE + CO(2) + NADPH.  
CC -1- PATHWAY: HEXOSE MONOPHOSPHATE SHUNT.  
CC -1- SIMILARITY: TO OTHER PROKARYOTIC AND EUKARYOTIC 6-PHOSPHOGLUCONATE  
CC DEHYDROGENASES.  
DR EMBL; Z73041; E243686; -.  
KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE; PENTOSE SHUNT; NADP.  
SQ SEQUENCE 492 AA; 53922 MW; 28FB6BB4 CRC32;

Query Match 63.9%; Score 46; DB 1; Length 492;  
Best Local Similarity 75.0%; Pred. No. 6.47e+00;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 101 idggnsfhf 108  
:|||||  
QY 2 LDGKNKF 9

Search completed: Tue Jun 10 11:24:05 1997  
Job time : 10 secs.

\*\*\*\*\*  
W E S L E H  
\*\*\*\*\*  
(TM)

\*\*\*\*\*  
Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
\*\*\*\*\*

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:26:12 1997; MasPar time 1.83 Seconds  
Tabular output not generated. 53.515 Million cell updates/sec

Title: >US-08-231-565A-34  
Description: (1-9) from US08231565A.pep  
Perfect Score: 77  
Sequence: 1 VLKRCLLHL 9

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 17.244; Variance 54.125; scale 0.319

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	77	100.0	9 15 R84202	gp100 melanoma antige	8.38e-02
2	77	100.0	10 15 R82145	Melanoma-specific mut	8.38e-02
3	77	100.0	11 15 R82192	Melanoma-specific mut	8.38e-02
4	77	100.0	11 15 R82191	Melanoma-specific mut	8.38e-02
5	77	100.0	661 15 R84854	MART-1 melanoma antig	8.38e-02
6	77	100.0	661 15 R84855	MART-1 melanoma antig	8.38e-02
7	77	100.0	661 14 R78646	Melanoma associated a	8.38e-02
8	60	77.9	9 15 R82109	Melanoma-specific mut	6.51e+00
9	53	68.8	751 7 R33424	G6PD.	3.59e+01
10	51	66.2	164 13 R66945	Deduced sequence of h	5.78e+01
11	50	64.9	394 2 P70434	Amino acids 71-464 of	7.32e+01
12	50	64.9	394 1 R04834	N-myc protein	7.32e+01
13	50	64.9	501 18 W03758	Mullerian inhibiting	7.32e+01
14	50	64.9	501 8 R41923	MISR4.	7.32e+01
15	50	64.9	3135 11 R57474	P. falciparum transmi	7.32e+01
16	49	63.6	220 11 R21500	Azurophil derived bac	9.25e+01
17	49	63.6	220 11 R60124	Azurophil-derived bac	9.25e+01
18	48	62.3	966 14 R76065	Yeast MSH2 protein.	1.17e+02
19	48	62.3	3210 1 P81771	Deduced sequence enco	1.17e+02
20	47	61.0	63 9 R55643	GGF segment A'.	1.47e+02

21	47	61.0	63 13 R67226	Bovine glial cell gro	1.47e+02
22	47	61.0	63 9 R46902	GGF segment A'.	1.47e+02
23	47	61.0	63 17 R86694	Bovine glial growth f	1.47e+02
24	47	61.0	63 17 R96059	Bovine glial growth f	1.47e+02
25	47	61.0	137 6 R28535	Bovine GGF-II encoded	1.47e+02
26	47	61.0	137 17 R96085	Bovine glial growth f	1.47e+02
27	47	61.0	137 13 R67202	Bovine glial cell gro	1.47e+02
28	47	61.0	137 9 R46894	Putative bovine GGF-I	1.47e+02
29	47	61.0	137 9 R55691	Putative bovine GGF-I	1.47e+02
30	47	61.0	137 17 R86678	Bovine glial growth f	1.47e+02
31	47	61.0	206 9 R48895	GGF2BPPI.	1.47e+02
32	47	61.0	206 13 R67216	Putative bovine glial	1.47e+02
33	47	61.0	206 5 R28536	GGF2BPPI.CDS protein.	1.47e+02
34	47	61.0	206 9 R55688	GGF2BPPI.	1.47e+02
35	47	61.0	206 17 R86687	Bovine glial growth f	1.47e+02
36	47	61.0	206 17 R96082	Bovine glial growth f	1.47e+02
37	47	61.0	431 3 R13731	Human cyclin A.	1.47e+02
38	47	61.0	441 9 R43992	Bovine ER receptor su	1.47e+02
39	47	61.0	442 6 R30886	ETB receptor.	1.47e+02
40	47	61.0	691 2 R04711	Sequence of guinea pi	1.47e+02
41	47	61.0	1422 16 R82067	Hepatitis GB virus (H	1.47e+02
42	47	61.0	1463 12 R63575	Group 1 phospholipase	1.47e+02
43	47	61.0	1463 16 R85143	Bovine phospholipase	1.47e+02
44	47	61.0	4472 19 R97245	Virulence gene cluste	1.47e+02
45	46	59.7	29 3 P61720	Fragment of sak gene	1.85e+02

ALIGNMENTS

RESULT 1  
ID R84202 standard; Peptide; 9 AA.  
AC R84202;  
DT 25-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G9-19).  
KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN WO9529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 55; Page 131; 184pp; English.  
CC The immunogenic peptide is derived from cDNA25 (R84854), a  
CC melanoma antigen derivative of gp100 (see R84855). The  
CC peptide and its derivatives (see R84200-R84211) are used in  
CC medicaments (vaccines) for the treatment or prevention (by  
CC immunisation) of melanoma. Antibodies against melanoma-specific  
CC antigens and its immunogenic peptides may be used in the  
CC detection and isolation of the antigen from a sample, the  
CC detection of which is indicative of a disease state  
CC (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 100.0%; Score 77; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.38e-02;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vlkrcllhl 9  
| | | | |  
Qy 1 VLKRCLLHL 9

RESULT 2  
ID R82145 standard; peptide; 10 AA.

```

AC R82145;
DT 25-MAR-1996 (first entry)
DE Melanoma-specific mutant immunogen epitope 10mer peptide.
KW Melanoma; Immunogen; epitope; homologue; vaccine; immunotherapy;
KW cytotoxic T cell; lymphocyte; HLA-A2.
OS Homo sapiens.
PN W09522561-A2.
PD 24-AUG-1995.
PR 16-FEB-1995; US-197399.
PR 16-FEB-1994; US-197399.
PR 29-APR-1994; US-234784.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;
DR WPI; 95-302688/39.
PT Melanoma-specific immunogen comprises epitope(s) homologous with
PT pMel.17 - are highly potent stimulators of HLA-A2+CTL's useful in
PT adoptive immuno-therapy
PS Example 8; Page 52; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 comprises one
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-
CC R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines.
SQ Sequence 10 AA;

Query Match 100.0%; Score 77; DB 15; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.38e-02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vlkrcllhl 9
QY 1 VLKRCLLHL 9
|||||

RESULT 3
ID R82192 standard; peptide; 11 AA.
AC R82192;
DT 25-MAR-1996 (first entry)
DE Melanoma-specific mutant immunogen epitope 11mer peptide.
KW Melanoma; Immunogen; epitope; homologue; vaccine; immunotherapy;
KW cytotoxic T cell; lymphocyte; HLA-A2.
OS Homo sapiens.
PN W09522561-A2.
PD 24-AUG-1995.
PR 16-FEB-1995; US-197399.
PR 16-FEB-1994; US-197399.
PR 29-APR-1994; US-234784.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;
DR WPI; 95-302688/39.
PT Melanoma-specific immunogen comprises epitope(s) homologous with
PT pMel.17 - are highly potent stimulators of HLA-A2+CTL's useful in
PT adoptive immuno-therapy
PS Example 8; Page 53; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 comprises one
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-
CC R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines.
SQ Sequence 11 AA;

Query Match 100.0%; Score 77; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 8.38e-02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vlkrcllhl 9
QY 1 VLKRCLLHL 9
|||||

RESULT 3
ID R82192 standard; peptide; 11 AA.
AC R82192;
DT 25-MAR-1996 (first entry)
DE Melanoma-specific mutant immunogen epitope 11mer peptide.
KW Melanoma; Immunogen; epitope; homologue; vaccine; immunotherapy;
KW cytotoxic T cell; lymphocyte; HLA-A2.
OS Homo sapiens.
PN W09522561-A2.
PD 24-AUG-1995.
PR 16-FEB-1995; US-197399.
PR 16-FEB-1994; US-197399.
PR 29-APR-1994; US-234784.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;
DR WPI; 95-302688/39.
PT Melanoma-specific immunogen comprises epitope(s) homologous with
PT pMel.17 - are highly potent stimulators of HLA-A2+CTL's useful in
PT adoptive immuno-therapy
PS Example 8; Page 53; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 comprises one
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-
CC R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines.
SQ Sequence 11 AA;

Query Match 100.0%; Score 77; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 8.38e-02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vlkrcllhl 9
QY 1 VLKRCLLHL 9
|||||

```

```

QY 1 VLKRCLLHL 9
RESULT 4
ID R82191 standard; peptide; 11 AA.
AC R82191;
DT 25-MAR-1996 (first entry)
DE Melanoma-specific mutant immunogen epitope 11mer peptide.
KW Melanoma; Immunogen; epitope; homologue; vaccine; immunotherapy;
KW cytotoxic T cell; lymphocyte; HLA-A2.
OS Homo sapiens.
PN W09522561-A2.
PD 24-AUG-1995.
PR 16-FEB-1995; US-197399.
PR 16-FEB-1994; US-197399.
PR 29-APR-1994; US-234784.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;
DR WPI; 95-302688/39.
PT Melanoma-specific immunogen comprises epitope(s) homologous with
PT pMel.17 - are highly potent stimulators of HLA-A2+CTL's useful in
PT adoptive immuno-therapy
PS Example 8; Page 53; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 comprises one
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-
CC R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines.
SQ Sequence 11 AA;

Query Match 100.0%; Score 77; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 8.38e-02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 vlkrcllhl 11
QY 1 VLKRCLLHL 9
|||||

RESULT 5
ID R84854 standard; Protein; 661 AA.
AC R84854;
DT 08-MAY-1996 (revised)
DT 20-APR-1996 (first entry)
DE MART-1 melanoma antigen cDNA25.
KW cDNA25; MART-1; melanoma antigen recognised by T-cell;
KW gp100 antigen derivative; melanoma; metastatic melanoma;
KW tumour-associated antigen; immunogen; diagnosis; prognosis;
KW prophylaxis; therapy; vaccine.
OS Mammalian sp.
FH Key Location/Qualifiers
FT Peptide 457..466
FT /label= antigenic_peptide
FT /note= "see R84199"
FT W09529193-A2.
PN W09529193-A2.
PD 02-NOV-1995.
PR 21-APR-1995; US05063.
PR 22-APR-1994; US-231565.
PR 03-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
DR N-PSDB: T02716.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Claim 81; Fig 5A; 184pp; English.
CC cDNA2 is a melanoma antigen (MART-1) which is recognized by
CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-
CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source

```

CC of immunogenic peptides (see R84199) which are optionally modified  
 CC (see R84200-R84211) to enhance their binding to a MHC molecule and  
 CC used in medicaments, especially vaccines, for the treatment or  
 CC prevention (by immunisation) of melanoma. Antibodies against cDNA2  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of the antigen from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 77; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 8.38e-02;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 vlkrcllhl 12  
 |||||  
 QY 1 VLKRCLLHL 9

## RESULT 6

ID R84855 standard; Protein; 661 AA.  
 AC R84855;  
 DT 08-MAY-1996 (revised)  
 DE 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen gp100.  
 KW gp100; MART-1; melanoma antigen recognised by T-cell;  
 KW cDNA25 antigen derivative; melanocyte; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian sp.  
 FH Key Location/Qualifiers  
 FT Peptide 154..163  
 FT /label= G9-154\_immunogenic\_peptide  
 FT Peptide 208..217  
 FT /label= G9-209\_immunogenic\_peptide  
 FT /note= "see R84210"  
 FT Peptide 280..288  
 FT /label= G9-280\_immunogenic\_peptide  
 FT /note= "see R84208"  
 FT Peptide 457..266  
 FT /label= immunogenic\_peptide  
 FT Peptide 476..485  
 FT /label= immunogenic\_peptide  
 FT W09529193-A2.  
 PN 02-NOV-1995.  
 PD 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 DT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 7A; 184pp; English.  
 CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which  
 CC is recognized by T-lymphocytes, and is a derivative of the  
 CC melanoma-specific antigen cDNA25 (see R84854). gp100 is a source  
 CC of immunogenic peptides which are optionally modified to enhance  
 CC their binding to a MHC molecule, and used in medicaments,  
 CC especially vaccines, for the treatment or prevention (by  
 CC immunisation) of melanoma. Antibodies against cDNA2 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 77; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 8.38e-02;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 vlkrcllhl 12  
 |||||  
 QY 1 VLKRCLLHL 9

RESULT 7  
 ID R78646 standard; Protein; 661 AA.  
 AC R78646;  
 DT 22-JAN-1996 (first entry)  
 DE Melanoma associated antigen gp100.  
 KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 OS Homo sapiens.  
 PN EP-668350-A1.  
 PD 23-AUG-1995.  
 PF 14-FEB-1995; 200348.  
 PR 16-FEB-1994; EP-200337.  
 PR 21-DEC-1994; EP-203709.  
 PA (ALKU ) AKZO NOBEL NV.  
 PI Adema GJ, Figdor CG;  
 DR WPI; 95-284790/38.  
 DR N-PSDB: Q96055  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PT detection of tumours  
 PS Claim 1; Page 22-24; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC may be used in the production of vaccines. Nucleotide sequences  
 CC encoding the immunogenic peptides may be used as primers and probes  
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
 CC capable of binding to the melanoma associated antigen can be  
 CC cultured ex vivo and returned to melanoma particles, and when  
 CC radiolabelled, they may be used to identify tumour deposits.  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 77; DB 14; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 8.38e-02;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 vlkrcllhl 12  
 |||||  
 QY 1 VLKRCLLHL 9

## RESULT 8

ID R82109 standard; peptide; 9 AA.  
 AC R82109;  
 DT 25-MAR-1996 (first entry)  
 DE Melanoma-specific mutant immunogen epitope 9mer peptide.  
 KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
 KW cytotoxic T cell; lymphocyte; HLA-A2.  
 OS Homo sapiens.  
 PN W09522561-A2.  
 PD 24-AUG-1995.  
 PF 16-FEB-1995; U01991.  
 PR 16-FEB-1994; US-197399.  
 PR 29-APR-1994; US-234784.  
 PA (UYVI-) UNIV VIRGINIA PATENT FOUND.  
 PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
 DR WPI; 95-302688/39.  
 PT Melanoma-specific immunogen comprises epitope(s) homologous with  
 PT pmei.17 - are highly potent stimulators of HLA-A2-CTL's useful in  
 PT adoptive immuno-therapy  
 PS Example 8; Page 51; 148pp; English.  
 CC A melanoma-specific immunogen homologous with pmei-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
 CC R82108 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pmei-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.  
 SQ Sequence 9 AA;

Query Match 77.9%; Score 60; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 6.51e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db      3 vlkrcll 9
        |||||
QY      1 VLKRCLL 7

RESULT  9
ID      R33424 standard; Protein; 751 AA.
AC      R33424;
DT      27-JUL-1993 (first entry)
DE      G6PD.
KW      Polymerase chain reaction; glucose-6-phosphate dehydrogenase; PCR;
        E. coli; G6PD; drug; transformation; primer; amplify.
OS      Plasmodium falciparum.
PN      WO9306125-A.
PD      01-APR-1993.
PF      21-SEP-1992; U07807.
PR      20-SEP-1991; US-762137.
PI      (USSH ) US DEPT HEALTH & HUMAN SERVICE.
        Kaslow DC, Shahabuddin M;
DR      WPI; 93-117467/14.
DR      N-PSDB; Q33424.
PT      DNA segments encoding Plasmodium falciparum G6PD obtd. by PCR -
        used to express proteins and raise antibodies for diagnosis and
        PT treatment of malaria
PS      Claim 2; Fig 2; 36pp; English.
CC      This sequence represents Plasmodium falciparum glucose-6-phosphate
        dehydrogenase (G6PD). The DNA encoding this sequence was isolated
        using the primers given in Q38806-07. The amplified sequence was
        used in the production of transformed E. coli which produce a
        recombinant P. falciparum G6PD. These transformed cells can be used
        in a method of screening drugs for activity against P. falciparum
        CC G6PD.
SQ      Sequence 751 AA;

Query Match      68.8%; Score 53; DB 7; Length 751;
Best Local Similarity 100.0%; Pred. No. 3.59e+01;
Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      234 lkrcll 239
        |||||
QY      2 LKRCLL 7

RESULT  10
ID      R66945 standard; Protein; 164 AA.
AC      R66945;
DT      02-OCT-1995 (first entry)
DE      Deduced sequence of human methenyltetrahydrofolate synthetase.
KW      MTHFS; methenyltetrahydrofolate synthetase.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      Modified_site      11..15
FT      /label- phosphorylation site
FT      Modified_site      18..21
FT      /label- see above
FT      Modified_site      58..60
FT      /label- see above
FT      Peptide      95..100
FT      /label- sequenced peptide
FT      Binding_site      108..111
FT      /label- ATP consensus binding site
FT      Peptide      139..147
FT      /label- sequenced peptide
FT      Modified_site      152..157
FT      /label- phosphorylation site
PN      US9389516-A.
PD      14-FEB-1995.
PF      19-MAR-1993; Q33857.
PR      19-MAR-1993; US-Q33857.
PA      (UYMO-) UNIV MONTREAL.
PI      Bertrand R, Dayan A, Jolivet J;
DR      WPI; 93-090200/12.

DR      N-PSDB; Q84626.
PT      New cDNA probe for detecting meth:enyl:tetra:hydro:folate
        synthetase - to assess usefulness of 5-formyl:tetra:hydro:folate
        for treating cancer in conjunction with 5-fluoro:uracil
PT      Disclosure; Figure 2; 14pp; English.
PS      Human liver MTHFS was purified to homogeneity. Purified enzyme was
        chemically cleaved with cyanogen bromide (CnBr) and enzymatically
        digested by protease V8. A sequence of only six AAs (R66943) could
        be established from one of the CnBr fragments due to the small
        amounts of the purified protein left after purification and was
        used for MTHFS cDNA cloning. Degenerate oligos corresp. to a
        heptapeptide were used to amplify a 389 cDNA prep. from human
        CC liver poly(A)+ using a 3' RACE (R66944) protocol. The PCR product
        was subcloned and sequenced and found to contain nts corresp. to
        CC the primer sequence and to an ORF of 97 AAs. This PCR product was
        CC then used as a probe to screen the lambda DR2 human liver cDNA
        CC library and a clone of 867 bp was obtd. (see Q84626). The
        CC resulting protein (R66945) contains two sequenced peptides obtd.
        CC from purified MTHFS protein (see I and II FT); an ATP consensus
        CC binding site (see III FT) and 4 consensus putative phosphorylation
        CC sites (see IV FT).
SQ      Sequence 164 AA;

Query Match      66.2%; Score 51; DB 13; Length 164;
Best Local Similarity 85.7%; Pred. No. 5.78e+01;
Matches      6; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

Db      118 lkrclq 124
        |||||
QY      2 LKRCLLH 8

RESULT  11
ID      P70434 standard; protein; 394 AA.
AC      P70434;
DT      27-JAN-1991 (first entry)
DE      Amino acids 71-464 of the N-myc 1 protein C-terminal.
KW      N-myc 1 protein; cancer.
OS      Homo sapiens.
PN      WO8706940-A.
PD      19-NOV-1987.
PR      005-MAY-1987; U01046.
        06-MAY-1986; US-860276.
PA      (REGC) Univ of California.
PI      Souza LM
PI      Slamon DJ
DR      WPI; 87-334949/47.
DR      N-PSDB; N70693.
PT      Recombinant N-myc proto-oncogene - used for producing antibodies and
        PT identifying human cancers and for use in cancer therapy.
PS      Disclosure; Fig 1; 31pp; English.
CC      Human cancers, especially neuroblastoma, retinoblastoma and
        CC small cell lung cancer, can be identified and monitored
        CC immunologically by detecting the N-myc protein. This is
        CC useful in cancer diagnosis and cancer therapy. Polypeptides
        CC containing at least 6 amino acids of the sequence provided
        CC may be used to raise antibodies specific for the human
        CC N-myc protein.
SQ      Sequence 394 AA;

Query Match      64.9%; Score 50; DB 2; Length 394;
Best Local Similarity 83.3%; Pred. No. 7.32e+01;
Matches      5; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

Db      249 ilkrcl 254
        :|||
QY      1 VLKRL 6

RESULT  12
ID      R04834 standard; protein; 394 AA.
AC      R04834;
DT      10-MAR-1993 (revised)

```

24-SEP-1989 (first entry)  
 DE N-myc protein  
 KW Cancer; N-myc; therapy.  
 FH Key Location/Qualifiers  
 FT Peptide 26..69  
 FT /label-peptide I  
 FT /note="C-terminal region of N-myc"  
 FT Peptide 108...183  
 FT /label-peptide II  
 FT /note="C-terminal region of N-myc"  
 FT Peptide 108...1  
 FT /label-peptide III  
 FT /note="C-terminal region of N-myc"  
 FT Peptide 186...240  
 FT /label-peptide IV  
 FT /note="C-terminal region of N-myc"  
 FT Peptide 340...394  
 FT /label-peptide V  
 FT /note="C-terminal region of N-myc"  
 FT Peptide 241...339  
 FT /label-peptide VI  
 FT /note="C-terminal region of N-myc"  
 PN US4918162-A.  
 PD 17-APR-1990.  
 PE 05-OCT-1988; 253933.  
 PR 05-OCT-1988; US-253933.  
 PA (REGC) Univ of California.  
 PI Slamon DJ, Souza LM;  
 DR WPI; 90-156111/20.  
 DR N-PSDB; 004459.  
 PT New synthetic polypeptide(s) useful in cancer diagnosis-partic. of neural  
 FT and neuro-endocrine cancers, are in carboxy region of N-myc protein  
 PS Disclosure; P: English.  
 CC This N-myc protein is encoded by the clone N-myc 1, obtained on screening  
 CC of a human neuroblastoma cDNA library with probe pNB-1. Constitutive  
 CC polypeptides, within the C-terminal 394 residues of N-myc, are useful for  
 CC detecting the N-myc proteins. These are esp. useful for distinguishing  
 CC amongst morphologically similar tumour types and will eg enable early  
 CC detection of small cell lung cancers. Peptide VI is an antigenic or  
 CC haptenic polypeptide defining an epitopic site which is immunologically  
 CC cross reactive with both N-myc and E-myc. Peptides II and III are capable  
 CC of eliciting antibodies (Abs) for use eg in cancer therapy.  
 SQ Sequence 394 AA;

Query Match 64.9%; Score 50; DB 1; Length 394;  
 Best Local Similarity 83.3%; Pred. No. 7.32e+01;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 249 ilkrcl 254  
 :|||||  
 QY 1 VLKRL 6

RESULT 13  
 ID W03758 standard; Protein; 501 AA.  
 AC W03758;  
 DT 30-OCT-1996 (first entry)  
 DE Mullerian inhibiting substance receptor MISR4.  
 KW Mullerian inhibiting substance receptor; MISR; TGF-beta receptor;  
 KW transforming growth factor beta type I receptor; gene therapy;  
 KW wound healing; tumour treatment; rat inhibin.  
 OS Rattus sp.  
 FH Key Location/Qualifiers  
 FT Binding\_site 210..230  
 FT /label=ATP\_binding\_site  
 FT /note="corresponds to conserved GXGXXGVX(11-28)K  
 FT motif found in all serine/threonine kinases  
 FT and thought to form an ATP binding site"  
 PN US5538892-A.  
 PD 23-JUL-1996.  
 PF 18-MAR-1992; 853396.  
 PR 18-MAR-1992; US-853396.  
 PR 11-MAR-1993; US-029673.

PR 04-NOV-1993; US-149105.  
 PA (GEO ) GEN HOSPITAL CORP.  
 PA (UYDU-) UNIV DUKE..  
 PI Donahoe PK, Gustafson M, He W, Wang X;  
 DR WPI; 96-353830/35.  
 DR N-PSDB; T36068.  
 PT New isolated TGF-beta type I receptor DNA - used to develop prods  
 FT for diagnosis and therapy, e.g. for treating tumours or promoting  
 PT wound healing  
 PS Claim 6; Columns 47-50; 44pp; English.  
 CC Degenerate PCR primers were designed based on two highly conserved  
 CC regions within the cDNA encoding a murine activin receptor, human  
 CC and porcine TGF-beta type II receptor and the da1-1 receptor of  
 CC C.elegans. The primers (see T36072 and T36073) were used for  
 CC amplifying clones present in a 14.5 day foetal rat urogenital ridge  
 CC cDNA COS cell expression library. Four clones encoding portions of  
 CC four novel polypeptides (all putative serine/threonine kinases)  
 CC were obtained and inserts from these clones were used as probes to  
 CC isolate full-length cDNA sequences for each of the four TGF-beta  
 CC type I receptors. Each putative receptor comprises a hydrophobic  
 CC signal peptide of 19-23 residues, an extracellular, cysteine-rich,  
 CC hydrophilic, ligand-binding domain of 100-150 residues, a  
 CC intracellular serine/threonine kinase domain of 23-25 residues, an  
 CC residues and a short serine/threonine-rich tail. The present sequence  
 CC is that of MISR4 which appears to be a functional type I TGF-beta  
 CC receptor, i.e. it can restore a TGF-beta induced gene response in  
 CC non-responsive R1B mutant cells which lack an endogenous functional  
 CC type I receptor; it requires the presence of a functional type II  
 CC receptor to bind ligand and to signal; its kinase activity is required  
 CC for signalling and its kinase domain is capable of serine/threonine  
 CC autophosphorylation under in vitro conditions. Isolated DNA which  
 CC encodes the MISR4 sequence or which is able to hybridise to such DNA  
 CC under stringent conditions is claimed.  
 SQ Sequence 501 AA;

Query Match 64.9%; Score 50; DB 18; Length 501;  
 Best Local Similarity 83.3%; Pred. No. 7.32e+01;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 8 lrrcll 13  
 :|||||  
 QY 2 LKRCLL 7

RESULT 14  
 ID R41923 standard; Protein; 501 AA.  
 AC R41923;  
 DT 20-APR-1994 (first entry)  
 DE MISR4.  
 KW MIS; Mullerian Inhibitory Substance; receptor;  
 KW transforming growth factor; inhibin; BMP;  
 KW membrane serine/threonine kinase receptor;  
 KW bone morphogenesis protein.  
 OS Rattus rattus.  
 PN W09319177-A.  
 PD 30-SEP-1993.  
 PF 15-MAR-1993; U02387.  
 PR 18-MAR-1992; US-853396.  
 PR 11-MAR-1993; US-029673.  
 PA (GEO ) GEN HOSPITAL CORP.  
 PI Donahoe PK, Gustafson M, He WW;  
 DR WPI; 93-320743/40.  
 DR N-PSDB; 049766.  
 PT New receptors of the transforming growth factor-beta receptor and  
 PT family - comprising Mullerian Inhibitory Substance Receptors and  
 PT inhibin receptors  
 PS Claim 57; Fig 4; 59pp; English.  
 CC Misr1 (049763) is believed to encode an isoform of the rat  
 CC MIS receptor. Misr2A/misr2B (049764), misr3 (049765) and misr4  
 CC (049766) are believed to encode monomeric isoforms of the rat  
 CC inhibin receptor and/or BMP receptor.  
 SQ Sequence 501 AA;

Query Match 64.9%; Score 50; DB 8; Length 501;  
Best Local Similarity 83.3%; Pred.No. 7.32e+01;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 8 lrrcll 13  
QY 2 LKRCLL 7

RESULT 15  
ID R57474 standard; Protein; 3135 AA.  
AC R57474;  
DT 20-FEB-1995 (first entry)  
DE P. falciparum transmission blocking target antigen Pfs230.  
KW Protozoan; transmission blocking target antigen; Pfs230; malaria;  
vaccine.  
OS Plasmodium falciparum.  
PN W09417187.A.  
PD 04-AUG-1994.  
PF 18-JAN-1994; U00547.  
PR 29-JAN-1993; US-010409.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kaslow DC, Williamson KC;  
DR WPI; 94-264101/32.  
DR N-PSDB; Q67190.  
PT New Plasmodium falciparum transmission blocking target antigen -  
useful in antimalarial vaccines, also related DNA, expression  
PT vectors and transformed cells  
PS Claim 6; Page 24; 63pp; English.  
CC Pfs230 protein is administered to humans to prevent transmission of  
malaria by inducing a transmission blocking immune response. It  
can also be used to raise antibodies and for T and B cell epitope  
mapping. Pfs230 induces a high and long-lasting antibody titer and  
can be produced in large amounts at low cost.  
SQ Sequence 3135 AA;

Query Match 64.9%; Score 50; DB 11; Length 3135;  
Best Local Similarity 62.5%; Pred.No. 7.32e+01;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 210 lkrcimki 217  
QY 2 LKRCLLHL 9

Search completed: Tue Jun 10 11:26:22 1997  
Job time : 10 secs.

\*\*\*\*\*

W P E R L H (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:25:42 1997; Maspar time 2.70 Seconds  
Tabular output not generated. 94.923 Million cell updates/sec

Title: >US-08-231-565A-34  
Description: (1-9) from US08231565A.pap  
Perfect Score: 77  
Sequence: 1 VLKRLHL 9

Scoring table: PAM 150  
Gap 15

Searched: 89912 seqs, 28507787 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir50

1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 23.492; Variance 37.868; scale 0.620

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	77	100.0	661	13	glycoprotein gp100 p	4.92e-04
2	77	100.0	668	13	melanocyte-specific	4.92e-04
3	61	79.2	283	8	probable pahF protei	4.71e-01
4	61	79.2	483	8	salicylaldehyde dehy	4.71e-01
5	55	71.4	495	4	early E1B 55K protei	5.06e+00
6	55	71.4	496	4	early E1B 55K protei	5.06e+00
7	54	70.1	1691	4	nonstructural protei	7.41e+00
8	54	70.1	3898	4	genome polyprotein -	7.41e+00
9	53	68.8	224	10	hypothetical protein	1.08e+01
10	53	68.8	736	12	glucose-6-phosphate	1.08e+01
11	53	68.8	855	12	glucose-6-phosphate	1.08e+01
12	53	68.8	910	12	glucose-6-phosphate	1.08e+01
13	52	67.5	91	11	hypothetical protein	1.57e+01
14	52	67.5	205	16	Mad3 protein - mouse	1.57e+01
15	52	67.5	381	1	protein kinase (EC 2	1.57e+01
16	52	67.5	381	1	protein kinase (EC 2	1.57e+01
17	52	67.5	791	11	hypothetical protein	1.57e+01
18	51	66.2	102	2	proteinase inhibitor	2.27e+01
19	51	66.2	203	13	indoleacetate-lysine	2.27e+01
20	51	66.2	379	1	protein kinase (EC 2	2.27e+01
21	51	66.2	406	6	cyclin A - bovine	2.27e+01

22	51	66.2	472	4	WNA6F6	early E1B 52K protei	2.27e+01
23	51	66.2	551	12	S65293	probable membrane pr	2.27e+01
24	51	66.2	556	12	S49754	probable membrane pr	2.27e+01
25	51	66.2	1146	11	S46837	hypothetical protein	2.27e+01
26	51	66.2	1483	11	S30015	hypothetical protein	2.27e+01
27	50	64.9	101	10	164242	hypothetical protein	3.27e+01
28	50	64.9	170	6	A22937	transforming protein	3.27e+01
29	50	64.9	209	9	S31071	rpcf protein - Syncr	3.27e+01
30	50	64.9	394	13	159514	N-myc protein - huma	3.27e+01
31	50	64.9	456	2	TVHUMC	transforming protein	3.27e+01
32	50	64.9	456	2	S02249	transforming protein	3.27e+01
33	50	64.9	464	6	TVHUM2	transforming protein	3.27e+01
34	50	64.9	515	11	S56784	hypothetical protein	3.27e+01
35	50	64.9	713	11	A56439	peroxisomal import p	3.27e+01
36	50	64.9	716	5	S07418	RNA-directed RNA pol	3.27e+01
37	50	64.9	716	1	P21VWS	RNA-directed RNA pol	3.27e+01
38	50	64.9	716	1	P21V68	RNA-directed RNA pol	3.27e+01
39	50	64.9	716	1	P21V61	RNA-directed RNA pol	3.27e+01
40	50	64.9	997	13	154390	Phospholipase C - hu	3.27e+01
41	50	64.9	1004	12	S53959	probable membrane pr	3.27e+01
42	50	64.9	1030	16	A42497	AE3 Cl-/HCO3- exchan	3.27e+01
43	50	64.9	1036	14	S62358	Inositol 1,4,5-trisp	3.27e+01
44	50	64.9	1227	7	B34911	band 3-related prote	3.27e+01
45	50	64.9	3135	12	A48584	transmission-blockin	3.27e+01

ALIGNMENTS

RESULT	1
ENTRY	A53668
TITLE	glycoprotein gp100 precursor, melanocyte lineage - human
ALTERNATE_NAMES	melanoma antigen 25
ORGANISM	#formal_name Homo sapiens #common_name man
DATE	07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 25-May-1996
ACCESSIONS	A53668; A55753
REFERENCE	A53668
#authors	Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G.
#journal	J. Biol. Chem. (1994) 269:20126-20133
#title	Molecular characterization of the melanocyte lineage-specific antigen gp100.
#accession	A53668
#molecule_type	mRNA
#residues	1-661
REFERENCE	A55753
#authors	Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Yanneli, J.R.; Adema, G.J.; Miki, T.; Rosenberg, S.A.
#journal	Proc. Natl. Acad. Sci. U.S.A. (1994) 91:6458-6462
#title	Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection.
#accession	A55753
#status	nucleic acid sequence not shown; not compared with conceptual translation
#molecule_type	mRNA
#residues	1-161, F', 163-661
KEYWORDS	glycoprotein
SUMMARY	#length 661 #molecular-weight 70255 #checksum 5487

Query Match 100.0%; Score 77; DB 13; Length 661;  
Best Local Similarity 100.0%; Pred. No. 4.92e-04;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 vlkrc1hl 12  
QY 1 VLKRLHL 9

RESULT 2  
ENTRY  
TITLE

A41234 #type complete  
melanocyte-specific protein Pmel-17 precursor - human



```

ORGANISM #formal_name Homo sapiens #common_name man
DATE 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS A41234
REFERENCE A41234
#authors Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
          Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
          Kobayashi, Y.; Kim, K.K.
#journal Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title A melanocyte-specific gene, Pmel 17, maps near the silver
        coat color locus on mouse chromosome 10 and is in a
        syntenic region on human chromosome 12.
#cross-references MUID:92021023
#accession A41234
#status preliminary
#molecule_type mRNA
#residues 1-668 #label KWO
#cross-references GB:M77348
SUMMARY #length 668 #molecular-weight 70932 #checksum 6409

Query Match 100.0%; Score 77; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 4.92e-04;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 vlkrcllhl 12
   |||||
QY 1 VLKRCLLHL 9

RESULT 3
ENTRY #type fragment
TITLE Probable pahF protein - Pseudomonas putida (strain OUS82)
ORGANISM #formal_name Pseudomonas putida
DATE 05-May-1995 #sequence_revision 05-May-1995 #text_change
ACCESSIONS F55217
REFERENCE F55217
#authors Takizawa, N.; Kaida, N.; Torigoe, S.; Moritani, T.; Sawada,
          T.; Satoh, S.; Kiyohara, H.
#journal J. Bacteriol. (1994) 176:2444-2449
#title Identification and characterization of genes encoding
        polycyclic aromatic hydrocarbon dioxygenase and polycyclic
        aromatic hydrocarbon dihydrodiol dehydrogenase in
        Pseudomonas putida OUS82.
#accession F55217
#status preliminary
#molecule_type DNA
#residues 1-283 #label TAK
#cross-references GB:D16629
SUMMARY #length 283 #checksum 4014

Query Match 79.2%; Score 61; DB 8; Length 283;
Best Local Similarity 87.5%; Pred. No. 4.71e-01;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 244 lkrcillel 251
   |||||
QY 2 LKRCLLHL 9

RESULT 4
ENTRY #type complete
TITLE salicylaldehyde dehydrogenase (EC 1.2.1.-) - Pseudomonas sp.
        (strain Cl8)
ORGANISM #formal_name Pseudomonas sp.
DATE 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change
ACCESSIONS E49343; S27636
REFERENCE E49343
#authors Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.
#journal J. Bacteriol. (1993) 175:6890-6901
#title Metabolism of dibenzothioephene and naphthalene in Pseudomonas

```

---

```

strains: complete DNA sequence of an upper naphthalene
catabolic pathway.
#accession E49343
#status preliminary
#molecule_type DNA
#residues 1-483 #label DEN
#cross-references GB:M60405
GENETICS doxR
#gene oxidoreductase
KEYWORDS #length 483 #molecular-weight 51994 #checksum 6328
SUMMARY Query Match 79.2%; Score 61; DB 8; Length 483;
        Best Local Similarity 87.5%; Pred. No. 4.71e-01;
        Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 244 lkrcillel 251
   |||||
QY 2 LKRCLLHL 9

RESULT 5
ENTRY #type complete
TITLE early E1B 55K protein I - human adenovirus 2
ORGANISM #formal_name Mastadenovirus h2 #common_name human adenovirus
        2
#note host Homo sapiens (man)
DATE 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change
ACCESSIONS B03809; A03809
REFERENCE B03809; A03809
#authors Gerasas, T.R.; Sciaky, D.; Gelinas, R.E.; Bing-Dong, J.;
          Yen, C.E.; Kelly, M.M.; Bullock, P.A.; Parsons, B.L.;
          O'Neill, K.E.; Roberts, R.J.
#journal J. Biol. Chem. (1982) 257:13475-13491
#title Nucleotide sequences from the adenovirus-2 genome.
#cross-references MUID:83056843
#accession B03809
#molecule_type DNA
#residues 1-495 #label GIN
GENETICS #map_position 5.5-9.6
CLASSIFICATION #superfamily adenovirus early E1B protein I
        early protein
KEYWORDS #length 495 #molecular-weight 54895 #checksum 1799

Query Match 71.4%; Score 55; DB 4; Length 495;
Best Local Similarity 66.7%; Pred. No. 5.06e+00;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 388 iltrcsllhl 396
   |||||
QY 1 VLKRCLLHL 9

RESULT 6
ENTRY #type complete
TITLE early E1B 55K protein I - human adenovirus 5
ORGANISM #formal_name Mastadenovirus h5 #common_name human adenovirus
        5
#note host Homo sapiens (man)
DATE 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change
ACCESSIONS A03809
REFERENCE A03809
#authors Bos, J.L.; Polder, L.J.; Bernards, R.; Schrier, P.I.; van den
          Elsen, P.J.; van der Eb, A.J.; van Ormondt, H.
#journal Cell (1981) 27:121-131
#title The 2.2 kb E1B mRNA of human Ad12 and Ad5 codes for two tumor
        antigens starting at different AUG triplets.
#cross-references MUID:82115327
#accession A03809
#molecule_type mRNA

```

```

##residues      1-496  ##label BOS
GENETICS
#map_position 5.5-9.6
CLASSIFICATION #superfamily adenovirus early E1B protein I
KEYWORDS
#length 496 #molecular-weight 54999 #checksum 2702
SUMMARY

Query Match      71.4%; Score 55; DB 4; Length 496;
Best Local Similarity 66.7%; Pred. No. 5.06e+00;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 389 iltrcs1hl 397
QY 1 VLKRCLLHL 9

RESULT 7
ENTRY
TITLE nonstructural protein - hepatitis E virus (strain Mexico)
ORGANISM #formal_name hepatitis E virus
DATE 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change
08-Dec-1994
ACCESSIONS A44212; B48547
REFERENCE
#authors Huang, C.C.; Nguyen, D.; Fernandez, J.; Yun, K.Y.; Fry, K.E.;
Bradley, D.W.; Tam, A.W.; Reyes, G.R.
#journal Virology (1992) 191:550-558
#title Molecular cloning and sequencing of the Mexico isolate of
hepatitis E virus (HEV).
#accession A44212
##molecule_type genomic RNA
##residues 1-1691 ##label HUA
##cross-references GB:M74506
REFERENCE
#authors Fry, K.E.; Tam, A.W.; Smith, M.M.; Kim, J.P.; Luk, K.C.;
Young, L.M.; Piatak, M.; Feldman, R.A.; Yun, K.Y.; Purdy,
M.A.; McCaustland, K.A.; Bradley, D.W.; Reyes, G.R.
#journal Virus Genes (1992) 6:173-185
#title Hepatitis E virus (HEV): strain variation in the
nonstructural gene region encoding consensus motifs for an
RNA-dependent RNA polymerase and an ATP/GTP binding site.
#cross-references M01D:92271462
#accession B48547
##molecule_type genomic RNA
##residues 965-1691 ##label FRY
##cross-references NCBI:104576; NCBI:104578
##note sequence extracted from NCBI backbone
CLASSIFICATION #superfamily hepatitis E virus nonstructural protein
KEYWORDS ATP binding; GTP binding; nonstructural protein
FEATURE
973-980 #region nucleotide-binding motif A (P-loop)\
979 #binding_site ATP/GTP (Lys) #status predicted
SUMMARY #length 1691 #molecular-weight 185223 #checksum 835

Query Match      70.1%; Score 54; DB 4; Length 1691;
Best Local Similarity 75.0%; Pred. No. 7.41e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 102 vlhrcflh 109
QY 1 VLKRCLLH 8

RESULT 8
ENTRY
TITLE genome polyprotein - bovine viral diarrhea virus (strain
SD-1)
ORGANISM #formal_name bovine viral diarrhea virus, BVDV
DATE 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change
31-May-1996
ACCESSIONS A44217
REFERENCE
#authors Deng, R.; Brock, K.V.

#journal Virology (1992) 191:867-879
#title Molecular cloning and nucleotide sequence of a pestivirus
genome, noncytopathic bovine viral diarrhea virus strain
SD-1.
#accession A44217
##molecule_type genomic RNA
##residues 1-3898 ##label DEN
##cross-references GB:M96751
##note this polyprotein may be cleaved into several mature
proteins, including p20 protein, p14 protein, gp48
protein, gp25 protein, gp53 protein, p54 protein, p80
protein, p10 protein, p38 protein, and p15 protein;
the cleavage sites are not reported
CLASSIFICATION #superfamily pestivirus genome polyprotein
KEYWORDS glycoprotein; polyprotein; RNA binding; zinc finger
FEATURE
253-265 #region hydrophobic\
547-562 #region hydrophobic\
556-670 #region hydrophobic\
675-694 #region hydrophobic\
1031-1046 #region hydrophobic\
1074-1099 #region hydrophobic\
1149-1164 #region hydrophobic\
1217-1238 #region hydrophobic\
1252-1269 #region hydrophobic\
1271-1292 #region hydrophobic\
1293-1304 #region hydrophobic\
1357-1373 #region hydrophobic\
1484-1512 #region zinc finger\
2562-2582 #region hydrophobic\
272-281,296,335,
365,370,413,487,
597,809,878,922,
990,1357,1419,1713,
2134,2217,2494,
2682,2751,2891,
2988,3688,3777,
3793 #binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 3898 #molecular-weight 437805 #checksum 8806

Query Match      70.1%; Score 54; DB 4; Length 3898;
Best Local Similarity 75.0%; Pred. No. 7.41e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1230 lirc1hl 1237
QY 2 LKRCLLHL 9

RESULT 9
ENTRY
TITLE hypothetical protein MJ0317 - Methanococcus jannaschii
ORGANISM #formal_name Methanococcus jannaschii
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
13-Sep-1996
ACCESSIONS F64339
REFERENCE
#authors Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,
R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann,
J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,
J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
Science (1996) 273:1058-1073
#journal Complete genome sequence of the methanogenic archaeon,
Methanococcus jannaschii.
#title
#accession F64339
#status preliminary; nucleic acid sequence not shown;

```

```

translation not shown
##molecule_type DNA
##residues 1-224 ##label BUL
##cross-references GB:L77117; TIGR:MT0317; CDS_PID:g1510408
GENETICS
#map_position REV296404-295730
#start_codon TTG
SUMMARY #length 224 #molecular-weight 26112 #checksum 3736

Query Match      68.8%; Score 53; DB 10; Length 224;
Best Local Similarity 100.0%; Pred. No. 1.08e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 172 lkrcll 177
|||||
QY 2 LKRCLL 7

RESULT 10
ENTRY #type complete
TITLE glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) -
ORGANISM Plasmodium falciparum
DATE #formal_name Plasmodium falciparum
17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change
08-May-1993
ACCESSIONS S27830
REFERENCE Shahabuddin, M.; Rawlings, D.J.; Kaslow, D.C.
#authors submitted to the EMBL Data Library, February 1992
#description Plasmodium falciparum glucose-6-phosphate dehydrogenase
deduced amino acid sequence.
#accession S27830
##molecule_type mRNA
##residues 1-736 ##label SHA
##cross-references EMBL:M80655
KEYWORDS oxidoreductase
SUMMARY #length 736 #molecular-weight 86358 #checksum 3092

Query Match      68.8%; Score 53; DB 12; Length 736;
Best Local Similarity 100.0%; Pred. No. 1.08e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 219 lkrcll 224
|||||
QY 2 LKRCLL 7

RESULT 11
ENTRY #type complete
TITLE glucose-6-phosphate dehydrogenase - Plasmodium falciparum
ORGANISM #formal_name Plasmodium falciparum
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
20-Feb-1995
ACCESSIONS S47533
REFERENCE Shahabuddin, M.; Rawlings, D.J.; Kaslow, D.C.
#authors Biochim. Biophys. Acta (1994) 1219:191-194
#journal A novel glucose-6-phosphate dehydrogenase in Plasmodium
#title falciparum: cDNA and primary protein structure.
#accession S47533
##status preliminary
##molecule_type mRNA
##residues 1-855 ##label SHA
##cross-references EMBL:M80655
SUMMARY #length 855 #molecular-weight 100422 #checksum 4125

Query Match      68.8%; Score 53; DB 12; Length 855;
Best Local Similarity 100.0%; Pred. No. 1.08e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 338 lkrcll 343
|||||
QY 2 LKRCLL 7

translation not shown
##molecule_type DNA
##residues 1-224 ##label BUL
##cross-references GB:L77117; TIGR:MT0317; CDS_PID:g1510408
GENETICS
#map_position REV296404-295730
#start_codon TTG
SUMMARY #length 224 #molecular-weight 26112 #checksum 3736

Query Match      68.8%; Score 53; DB 10; Length 224;
Best Local Similarity 100.0%; Pred. No. 1.08e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 172 lkrcll 177
|||||
QY 2 LKRCLL 7

RESULT 12
ENTRY #type complete
TITLE glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) -
ORGANISM Plasmodium falciparum
DATE #formal_name Plasmodium falciparum
06-Mar-1994 #sequence_revision 10-Nov-1995 #text_change
10-Nov-1995
ACCESSIONS S40259
REFERENCE O'Brien, E.; Kurdi-Haidar, B.; Manachiwanawin, W.; Carvajal,
J.; Vulliamy, T.; Cappadoro, M.; Mason, P.J.; Luzzatto, L.
#authors submitted to the EMBL Data Library, September 1993
#description Cloning of the glucose-6-phosphate dehydrogenase gene from
Plasmodium falciparum.
#accession S40259
##molecule_type DNA
##residues 1-910 ##label OBR
##cross-references EMBL:X74988
KEYWORDS oxidoreductase
SUMMARY #length 910 #molecular-weight 106986 #checksum 601

Query Match      68.8%; Score 53; DB 12; Length 910;
Best Local Similarity 100.0%; Pred. No. 1.08e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 395 lkrcll 400
|||||
QY 2 LKRCLL 7

RESULT 13
ENTRY #type complete
TITLE hypothetical protein - Common tobacco
ORGANISM #formal_name Nicotiana tabacum #common_name common tobacco
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
06-Jan-1995
ACCESSIONS S37190
REFERENCE Phillips, J.; Conrad, U.
#authors submitted to the EMBL Data Library, September 1993
#description Genomic sequences from Nicotiana tabacum homologous to the
maize transcriptional activator gene Vpl.
#accession S37190
##status preliminary
##molecule_type DNA
##residues 1-91 ##label PHI
##cross-references EMBL:Z26310
SUMMARY #length 91 #molecular-weight 10535 #checksum 5492

Query Match      67.5%; Score 52; DB 11; Length 91;
Best Local Similarity 62.5%; Pred. No. 1.57e+01;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 19 lkrcll 26
:|||||
QY 1 VLKRCLLH 8

RESULT 14
ENTRY #type complete
TITLE Mad3 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 23-Aug-1996 #sequence_revision 23-Aug-1996 #text_change
23-Aug-1996
ACCESSIONS S60005
REFERENCE Hurlin, P.J.; Queva, C.; Koskinen, P.J.; Steingrimsen, E.;
#authors Ayer, D.E.; Copeland, N.G.; Jenkins, N.A.; Eisenman, R.N.
#journal EMBO J. (1995) 14:5646-5659
#title Mad3 and Mad4: novel Max-interacting transcriptional
repressors that suppress c-myc dependent transformation and

```

#accession S60005 are expressed during neural and epidermal differentiation.

##status preliminary

##residues 1-205 ##label HUR

SUMMARY #length 205 #molecular-weight 23505 #checksum 4287

Query Match 67.5%; Score 52; DB 16; Length 205;

Best Local Similarity 75.0%; Pred. NO. 1.57e+01;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 73 lkrclql 80

|||||

Qy 2 LKRCLLHL 9

RESULT 15

ENTRY

TITLE A60669 #type complete

protein kinase (EC 2.7.1.37), cAMP-dependent, type I-beta

regulatory chain - rat

ORGANISM #formal\_name Rattus norvegicus #common\_name Norway rat

DATE 28-Apr-1993 #sequence\_revision 27-Jun-1994 #text\_change

10-Feb-1995

ACCESSIONS A60669

REFERENCE A60669

#authors Massa, J.S.; Fellows, R.E.; Maurer, R.A.

#journal Mol. Reprod. Dev. (1990) 26:129-133

#title Rat Ribeta isoform of type I regulatory subunit of cyclic

adenosine monophosphate-dependent protein kinase: cDNA

sequence analysis, mRNA tissue specificity, and rat/mouse

difference in expression in testis.

#accession A60669

##molecule\_type mRNA

##residues 1-381 #label MAS

COMMENT The inactive form of the enzyme is composed of two regulatory chains and two catalytic chains. Activation by cAMP produces two active catalytic monomers and a regulatory dimer that binds four cAMP molecules.

COMMENT Four types of regulatory chains are found: I-alpha, I-beta, II-alpha, and II-beta. Their expression varies among tissues and is in some cases constitutive and in others inducible.

COMMENT Type I regulatory chains contain a high-affinity binding site for Mg-ATP.

COMMENT The pseudophosphorylation site binds to the substrate-binding region of the catalytic chain but is not phosphorylated. The physiological significance of phosphorylations by other kinases is unclear.

CLASSIFICATION #superfamily cAMP-dependent protein kinase regulatory chain; cAMP receptor protein cyclic nucleotide-binding domain homology

acetylated amino end; ATP binding; cAMP binding; duplication; heterotetramer; homodimer; phosphotransferase

KEYWORDS

FEATURE

2-381

#product protein kinase, cAMP-dependent, type I-beta

regulatory chain #status predicted #label MAT\

2-136 #domain protein interaction #label DIM\

96-100 #region pseudophosphorylation motif\

137-254 #domain cAMP receptor protein cyclic nucleotide-binding domain homology #label CA1\

255-381 #domain cAMP receptor protein cyclic nucleotide-binding domain homology #label CA2\

2 #modified\_site acetylated amino end (Ala) (in mature form) #status predicted\

18 #disulfide\_bonds\_interchain (to 39) #status predicted\

39 #disulfide\_bonds\_interchain (to 18) #status predicted\

202,211 #binding\_site cAMP (Glu, Arg) #status predicted\

326,335 #binding\_site cAMP (Glu, Arg) #status predicted\

SUMMARY #length 381 #molecular-weight 43282 #checksum 6234

Query Match 67.5%; Score 52; DB 1; Length 381;

Best Local Similarity 66.7%; Pred. NO. 1.57e+01;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 30 vlkecihl 38

Qy 1 VLKRCLLHL 9

Search completed: Tue Jun 10 11:25:54 1997

Job time : 12 secs.

\*\*\*\*\*  
W P S R L A  
(TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:25:17 1997; MasPar time 2.09 Seconds  
Tabular output not generated. 91.511 Million cell updates/sec

Title: >US-08-231-565A-34  
Description: (1-9) from US08231565A.pep  
Perfect Score: 77  
Sequence: 1 VLKRCLLHL 9  
Scoring table: PAM 150  
Gap 15  
Searched: 59021 seqs, 21210388 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 24.565; Variance 31.412; scale 0.782

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.	
1	77	100.0	668	7	PM17_HUMAN	PMEL 17 PROTEIN PRECU	1.28e-05
2	55	71.4	495	3	E1BL_ADE02	E1B PROTEIN, LARGE T-	9.86e-01
3	55	71.4	496	3	E1BL_ADE05	E1B PROTEIN, LARGE T-	9.86e-01
4	54	70.1	1691	7	POLN_HEVME	NON-STRUCTURAL POLYPR	1.56e+00
5	54	70.1	3898	7	POLG_BVDVS	GENOME POLYPROTEIN	1.56e+00
6	52	67.5	380	5	KAP1_MOUSE	CAMP-DEPENDENT PROTEI	3.85e+00
7	51	66.2	202	4	FTHC_HUMAN	5-FORMYL-TETRAHYDROFOL	6.00e+00
8	51	66.2	380	5	KAP1_HUMAN	CAMP-DEPENDENT PROTEI	6.00e+00
9	51	66.2	406	2	CG2A_BOVIN	G2/MITOTIC-SPECIFIC C	6.00e+00
10	51	66.2	472	3	E1BL_ADE41	E1B PROTEIN, LARGE T-	6.00e+00
11	51	66.2	1146	11	YHC3_YEAST	HYPOTHETICAL 130.0 KD	6.00e+00
12	51	66.2	1483	11	YKB0_YEAST	HYPOTHETICAL 167.8 KD	6.00e+00
13	50	64.9	101	10	Y388_MYCCE	HYPOTHETICAL PROTEIN	9.27e+00
14	50	64.9	209	8	RPCF_SYNPY	BILIN BIOSYNTHESIS PR	9.27e+00
15	50	64.9	464	6	MYCN_HUMAN	N-MYC PROTO-ONCOGENE	9.27e+00
16	50	64.9	501	5	KIR4_RAT	SERINE/THREONINE-PROT	9.27e+00
17	50	64.9	515	11	YJB3_YEAST	HYPOTHETICAL 59.5 KD	9.27e+00
18	50	64.9	716	8	RRP2_IAZH3	RNA-DIRECTED RNA POLY	9.27e+00
19	50	64.9	716	8	RRP2_IAZTF	RNA-DIRECTED RNA POLY	9.27e+00
20	50	64.9	716	8	RRP2_IAZLO	RNA-DIRECTED RNA POLY	9.27e+00
21	50	64.9	716	8	RRP2_IAZ11	RNA-DIRECTED RNA POLY	9.27e+00
22	50	64.9	716	8	RRP2_IAYKM	RNA-DIRECTED RNA POLY	9.27e+00

23	50	64.9	716	8	RRP2_IAZH2	RNA-DIRECTED RNA POLY	9.27e+00
24	50	64.9	716	8	RRP2_IAY17	RNA-DIRECTED RNA POLY	9.27e+00
25	50	64.9	716	8	RRP2_IAIE7	RNA-DIRECTED RNA POLY	9.27e+00
26	50	64.9	716	8	RRP2_IAPPR	RNA-DIRECTED RNA POLY	9.27e+00
27	50	64.9	716	8	RRP2_IAGU2	RNA-DIRECTED RNA POLY	9.27e+00
28	50	64.9	716	8	RRP2_IAGU2	RNA-DIRECTED RNA POLY	9.27e+00
29	50	64.9	716	8	RRP2_IAGU2	RNA-DIRECTED RNA POLY	9.27e+00
30	50	64.9	716	8	RRP2_IAGU2	RNA-DIRECTED RNA POLY	9.27e+00
31	50	64.9	716	8	RRP2_IAGU2	RNA-DIRECTED RNA POLY	9.27e+00
32	50	64.9	716	8	RRP2_IAPUD	RNA-DIRECTED RNA POLY	9.27e+00
33	50	64.9	716	8	RRP2_IAPUD	RNA-DIRECTED RNA POLY	9.27e+00
34	50	64.9	716	8	RRP2_IAPUE	RNA-DIRECTED RNA POLY	9.27e+00
35	50	64.9	716	8	RRP2_IAPUE	RNA-DIRECTED RNA POLY	9.27e+00
36	50	64.9	1004	11	YGI1_YEAST	HYPOTHETICAL 113.9 KD	9.27e+00
37	50	64.9	1227	1	B3A3_RAT	ANION EXCHANGE PROTEI	9.27e+00
38	50	64.9	3135	9	S230_PLAFO	TRANSMISSION-BLOCKING	9.27e+00
39	49	63.6	221	3	C7FB_CLOAB	BUTYRATE-ACETOACETATE	1.42e+01
40	49	63.6	350	11	YCYA_EUGGR	HYPOTHETICAL 40.9 KD	1.42e+01
41	49	63.6	393	6	METK_ASCIM	S-ADENOSYLMETHIONINE	1.42e+01
42	49	63.6	863	11	YK44_YEAST	PUTATIVE 101.8 KD TRA	1.42e+01
43	49	63.6	1095	6	NKC2_MOUSE	BUMETANIDE-SENSITIVE	1.42e+01
44	49	63.6	1099	6	NKC2_RABIT	BUMETANIDE-SENSITIVE	1.42e+01
45	49	63.6	1432	9	SKI3_YEAST	SUPERKILLER 3 PROTEIN	1.42e+01

ALIGNMENTS

RESULT 1	PM17_HUMAN	STANDARD;	PRT;	668 AA.
AC	P40967;			
DT	01-FEB-1995 (REL. 31, CREATED)			
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	PMEL 17 PROTEIN PRECURSOR.			
GN	PMEL17.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 92021023.			
RA	KWON B.S., CHINTAMANANI C., KOZAK C.A., COPELAND N.G.,			
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,			
RA	KIM K.-K.			
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).			
CC	-1- FUNCTION: COULD BE A MELANOGENIC ENZYME.			
CC	-1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).			
DR	EMBL; M77348; GI90106; -.			
DR	MIM; 155550; -.			
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.			
FT	SIGNAL	1	23	POTENTIAL.
FT	CHAIN	24	668	PMEL 17 PROTEIN.
FT	TRANSMEM	575	595	POTENTIAL.
FT	TRANSMEM	603	623	POTENTIAL.
FT	DOMAIN	217	307	PKD.
FT	DOMAIN	315	444	10 X 13 AA TANDEM REPEATS.
FT	REPEAT	325	327	1.
FT	REPEAT	328	340	2.
FT	REPEAT	341	353	3.
FT	REPEAT	354	366	4.
FT	REPEAT	367	379	5.
FT	REPEAT	380	392	6.
FT	REPEAT	393	405	7.
FT	REPEAT	406	418	8.
FT	REPEAT	419	431	9.
FT	REPEAT	432	444	10.
FT	CARBOHYD	81	81	POTENTIAL.
FT	CARBOHYD	106	106	POTENTIAL.
FT	CARBOHYD	111	111	POTENTIAL.
FT	CARBOHYD	321	321	POTENTIAL.
FT	CARBOHYD	568	568	POTENTIAL.
SQ	SEQUENCE	668 AA;	70992 MW;	6E8E1AFO CRC32;

Query Match 100.0%; Score 77; DB 7; Length 668;  
Best Local Similarity 100.0%; Pred. No. 1.28e-05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 vlkrcllhl 12  
:|||||||  
QY 1 VLKRCLLHL 9

RESULT 2  
ID EIBL\_ADE02 STANDARD; PRT; 495 AA.  
AC P03244; 1986 (REL. 01, CREATED)  
DT 21-JUL-1986 (REL. 21, LAST SEQUENCE UPDATE)  
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)  
DE EIB PROTEIN, LARGE T-ANTIGEN.  
OS HUMAN ADENOVIRUS TYPE 2.  
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; ADENOVIRIDAE; MASTADENOVIRUSES.  
RN [1]  
RX MEDLINE; 83056843.  
RA GINGERAS T.R., SCIAXY D., GELINAS R.E., BING-DONG J., YEN C.E.,  
RA KELLY M.M., BULLOCK P.A., PARSONS B.D., O'NEILL K.E., ROBERTS R.J.,  
RL J. BIOL. CHEM. 257:13475-13491(1982).  
DR EMBL; J01917; G209820; -.  
DR PIR; B03809; Q1AD52.  
KW EARLY PROTEIN.  
SQ SEQUENCE 495 AA; 54908 MW; AFD2B5EE CRC32;

Query Match 71.4%; Score 55; DB 3; Length 495;  
Best Local Similarity 66.7%; Pred. No. 9.86e-01;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 388 iltrcsllhl 396  
:|||||||  
QY 1 VLKRCLLHL 9

RESULT 3  
ID EIBL\_ADE05 STANDARD; PRT; 496 AA.  
AC P03243;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)  
DE EIB PROTEIN, LARGE T-ANTIGEN.  
OS HUMAN ADENOVIRUS TYPE 5.  
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; ADENOVIRIDAE; MASTADENOVIRUSES.  
RN [1]  
RX MEDLINE; 82115327.  
RA BOS J.L., POLDER L.J., BERNARDS R., SCHRIER P.I., VAN DEN ELSEN P.J.,  
RA VAN DER EB A.J., VAN ORMONDT H.,  
RL CELL 27:121-131(1981).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 81165537.  
RA VAN ORMONDT H., MAAT J., VAN BEVEREN C.P.,  
RL GENE 11:299-309(1980).  
RN [3]  
RP COMPLETE GENOME.  
RX MEDLINE; 92087470.  
RA CHROBOCZEK J., BIEBER F., JACROT B.,  
RL VIROLOGY 186:280-285(1992).  
DR EMBL; M73260; -; NOT ANNOTATED\_CDS.  
DR EMBL; X02996; G58491; -.  
DR PIR; A03809; Q1AD55.  
KW EARLY PROTEIN.  
SQ SEQUENCE 496 AA; 54999 MW; 6AD66ECF CRC32;

Query Match 71.4%; Score 55; DB 3; Length 496;  
Best Local Similarity 66.7%; Pred. No. 9.86e-01;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 389 iltrcsllhl 397  
:|||||||  
QY 1 VLKRCLLHL 9

RESULT 4  
ID POLN\_HEVME STANDARD; PRT; 1691 AA.  
AC Q03495;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
DE NON-STRUCTURAL POLYPROTEIN (CONTAINS: RNA-DIRECTED RNA POLYMERASE  
(EC 2.7.7.48); HELICASE).  
OS HEPATITIS E VIRUS (STRAIN MEXICO) (HEV).  
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; CALCIVIRIDAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93079857.  
RA HUANG C.C., NGUYEN D., FERNANDEZ J., YUN K.Y., FRY K.E.,  
RA BRADLEY D.W., TAM A.W., REYES G.R.,  
RL VIROLOGY 191:550-558(1992).  
RN [2]  
RP SEQUENCE OF 965-1691 FROM N.A.  
RX MEDLINE; 92271462.  
RA FRY K.E., TAM A.W., SMITH M.M., KIM J.P., LUK K.C., YOUNG L.M.,  
RA PIATAK M., FELDMAN R.A., YUN K.Y., PURDY M.A., MCCAUSTLAND K.A.,  
RA BRADLEY D.W., REYES G.R.,  
RL VIRUS GENES 6:173-185(1992).  
CC -!- HEPATITIS E VIRUS IS THE MAJOR CAUSATIVE AGENT OF ENTERICALLY  
TRANSMITTED NON-A, NON-B HEPATITIS (ET-NANBH).  
DR EMBL; M74506; G330018; -.  
DR PIR; A44212; A44212.  
KW POLYPROTEIN; RNA-DIRECTED RNA POLYMERASE; HELICASE; ATP-BINDING.  
FT NP\_BIND 973 980 ATP (POTENTIAL).  
SQ SEQUENCE 1691 AA; 185224 MW; 1336ED66 CRC32;

Query Match 70.1%; Score 54; DB 7; Length 1691;  
Best Local Similarity 75.0%; Pred. No. 1.56e-00;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 102 vlhrcflh 109  
:|||||||  
QY 1 VLKRCLLHL 8

RESULT 5  
ID POLG\_BVDVS STANDARD; PRT; 3898 AA.  
AC Q01499;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE GENOME POLYPROTEIN.  
OS BOVINE VIRAL DIARRHEA VIRUS (STRAIN SD-1) (BVDV) (MUCOSAL DISEASE  
VIRUS).  
OC VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; FLAVIVIRIDAE;  
OC PESTIVIRUSES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93079889.  
RA DENG R., BROCK K.V.,  
RL VIROLOGY 191:867-869(1992).  
CC -!- FUNCTION: PESTIVIRUS P80 (P125) MAY BE A BIFUNCTIONAL PROTEIN  
WITH HELICASE AND PROTEASE ACTIVITY.  
CC -!- PTM: GP116 GIVES RISE TO GP62 AND GP53; GP62 IN TURN YIELDS GP48  
AND GP25.  
CC -!- SIMILARITY: TO THE HOG CHOLERA VIRUS GENOME POLYPROTEIN.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S31.  
DR EMBL; M96751; G289508; -.  
DR PIR; A44217; A44217.  
DR PROSITE; PS00531; RNASE\_T2.2.  
KW POLYPROTEIN; GLYCOPROTEIN; HELICASE; SERINE PROTEASE; HYDROLASE.  
FT CHAIN 1 7270 P20 (30KD).

```

FT CHAIN ?271 ?1063 GP116/GP62-GP53 (GLYCOPROTEIN).
FT CHAIN ? ? 3998
FT CHAIN ? ? 3998
FT DOMAIN 690 1658
FT ACT_SITE 1658 1658 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1695 1695 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 1752 1752 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 272 272 POTENTIAL.
FT CARBOHYD 281 281 POTENTIAL.
FT CARBOHYD 296 296 POTENTIAL.
FT CARBOHYD 335 335 POTENTIAL.
FT CARBOHYD 365 365 POTENTIAL.
FT CARBOHYD 370 370 POTENTIAL.
FT CARBOHYD 413 413 POTENTIAL.
FT CARBOHYD 487 487 POTENTIAL.
FT CARBOHYD 597 597 POTENTIAL.
FT CARBOHYD 809 809 POTENTIAL.
FT CARBOHYD 878 878 POTENTIAL.
FT CARBOHYD 922 922 POTENTIAL.
FT CARBOHYD 990 990 POTENTIAL.
FT CARBOHYD 1357 1357 POTENTIAL.
FT CARBOHYD 1419 1419 POTENTIAL.
FT CARBOHYD 1451 1451 POTENTIAL.
FT CARBOHYD 1713 1713 POTENTIAL.
FT CARBOHYD 2134 2134 POTENTIAL.
FT CARBOHYD 2217 2217 POTENTIAL.
FT CARBOHYD 2494 2494 POTENTIAL.
FT CARBOHYD 2682 2682 POTENTIAL.
FT CARBOHYD 2751 2751 POTENTIAL.
FT CARBOHYD 2891 2891 POTENTIAL.
FT CARBOHYD 2888 2888 POTENTIAL.
FT CARBOHYD 3688 3688 POTENTIAL.
FT CARBOHYD 3777 3777 POTENTIAL.
FT CARBOHYD 3793 3793 POTENTIAL.
SQ SEQUENCE 3898 AA; 437800 MW; A562145C CRC32;

Query Match 70.1%; Score 54; DB 7; Length 3898;
Best Local Similarity 75.0%; Pred. No. 1.56e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1230 lirc1hl 1237
| | | | |
QY 2 LKRCLLHL 9

RESULT 6 STANDARD; PRT; 380 AA.
ID KAP1_MOUSE
AC P12849;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CAMP-DEPENDENT PROTEIN KINASE TYPE I-BETA REGULATORY CHAIN.
GN PKAR1B.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE; 88234494.
RA CLEGG C.H., CADD G.G., MCKNIGHT G.S.;
RL PROC. NATL. ACAD. SCI. U.S.A. 85:3703-3707(1988).
CC -1- SUBUNIT: THE INACTIVE FORM OF THE ENZYME IS COMPOSED OF TWO
CC REGULATORY CHAINS AND TWO CATALYTIC CHAINS. ACTIVATION BY CAMP
CC PRODUCES TWO ACTIVE CATALYTIC MONOMERS AND A REGULATORY DIMER
CC THAT BINDS FOUR CAMP MOLECULES.
CC -1- PTM: THE PSEUDOPHOSPHORYLATION SITE BINDS TO THE SUBSTRATE-
CC BINDING REGION OF THE CATALYTIC CHAIN BUT IS NOT PHOSPHORYLATED.
CC THE PHYSIOLOGICAL SIGNIFICANCE OF PHOSPHORYLATIONS BY OTHER
CC KINASES IS UNCLEAR.
CC -1- TISSUE SPECIFICITY: FOUR TYPES OF REGULATORY CHAINS ARE FOUND:
CC I-ALPHA, I-BETA, II-ALPHA, AND II-BETA. THEIR EXPRESSION VARIES
CC AMONG TISSUES AND IS IN SOME CASES CONSTITUTIVE AND IN OTHERS

```

```

CC INDUCIBLE.
DR EMBL; M20473; G200365;
DR PIR; A30205; OKNSRL.
DR HSP; P00514; IAPK.
DR PROSITE; PS00888; CNMP_BINDING_1.
DR PROSITE; PS00889; CNMP_BINDING_2.
KW CAMP-BINDING; PHOSPHORYLATION; DUPLICATION; MULTIGENE FAMILY;
KW ACETYLATION.
FT INIT_MET 0 0
FT DOMAIN 1 135 DIMERIZATION AND PHOSPHORYLATION.
FT SITE 95 99 PSEUDOPHOSPHORYLATION MOTIF.
FT NP_BIND 136 253 CAMP.
FT NP_BIND 254 380 CAMP.
FT REPEAT 136 253
FT REPEAT 254 380
FT MOD_RES 1 1 ACETYLATION (PROBABLE).
FT DISULFID 17 17 INTERCHAIN (WITH 38 AND 17)
FT DISULFID 38 38 (BY SIMILARITY).
FT DISULFID 38 38 INTERCHAIN (WITH 38 AND 17)
FT MOD_RES 100 100 PHOSPHORYLATION (BY SIMILARITY).
FT BINDING 201 201 CAMP.
FT BINDING 210 210 CAMP.
FT BINDING 325 325 CAMP.
FT BINDING 334 334 CAMP.
SQ SEQUENCE 380 AA; 43093 MW; 9242886B CRC32;

Query Match 67.5%; Score 52; DB 5; Length 380;
Best Local Similarity 66.7%; Pred. No. 3.85e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 vlkecihl 37
| | | | |
QY 1 VLKRCLLHL 9

RESULT 7 STANDARD; PRT; 202 AA.
ID FTHC_HUMAN
AC P49914;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE 5-FORMYLTETRAHYDROFOLATE CYCLO-LIGASE (EC 6.3.3.2) (5,10-METHENYL-
DE TETRAHYDROFOLATE SYNTHETASE) (METHENYL-THF SYNTHETASE) (MTHFS).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 96096540.
RA DAYAN A., BERTRAND R., BEAUCHEMIN M., CHAHLA D., MAMO A.,
RA FILION M., SKUP D., MASSIE B., JOLIVET J.;
RL GENE 165:307-311(1995).
CC -1- CATALYTIC ACTIVITY: ATP + 5-FORMYLTETRAHYDROFOLATE = ADP +
CC ORTHOPHOSPHATE + 5,10-METHENYLTETRAHYDROFOLATE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: TO E.COLI YGFA.
DR EMBL; L38928; G886297;
KW LIGASE; FOLATE-BINDING; ACETYLATION.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
FT BINDING 17 17 FOLATE (BY SIMILARITY).
SQ SEQUENCE 202 AA; 23134 MW; 0040B23E CRC32;

Query Match 66.2%; Score 51; DB 4; Length 202;
Best Local Similarity 85.7%; Pred. No. 6.00e+00;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 156 lkrc1qlh 162
| | | | |
QY 2 LKRCLLHL 8

```

```

RESULT 8
ID KAP1_HUMAN STANDARD; PRT; 380 AA.
AC P31321;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CAMP-DEPENDENT PROTEIN KINASE TYPE I-BETA REGULATORY CHAIN.
GN PRKAR1B.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91207387.
RA SOLBERG R., TASKEN K., KEISERUD A., JAHNSEN T.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 176:166-172(1991).
RN [2]
RP REVISION TO 97-99.
RC TISSUE-TESTIS;
RX MEDLINE: 95010356.
RA SOLBERG R., TASKEN K., WEN W., COGHAN V.M., MEINKOTH J.L.,
RL SCOTT J.D., JAHNSEN T., TAYLOR S.S.;
EXP. CELL RES. 214:595-605(1994).
CC -1- SUBUNIT: THE INACTIVE FORM OF THE ENZYME IS COMPOSED OF TWO
REGULATORY CHAINS AND TWO CATALYTIC CHAINS. ACTIVATION BY CAMP
PRODUCES TWO ACTIVE CATALYTIC MONOMERS AND A REGULATORY DIMER
THAT BINDS FOUR CAMP MOLECULES.
CC -1- PTM: THE PSEUDOPHOSPHORYLATION SITE BINDS TO THE SUBSTRATE-
BINDING REGION OF THE CATALYTIC CHAIN BUT IS NOT PHOSPHORYLATED.
CC THE PHYSIOLOGICAL SIGNIFICANCE OF PHOSPHORYLATIONS BY OTHER
KINASES IS UNCLEAR.
CC -1- TISSUE SPECIFICITY: FOUR TYPES OF REGULATORY CHAINS ARE FOUND:
I-ALPHA, I-BETA, II-ALPHA, AND II-BETA. THEIR EXPRESSION VARIES
AMONG TISSUES AND IS IN SOME CASES CONSTITUTIVE AND IN OTHERS
INDUCIBLE.
EMBL: M65066; G307377; -.
DR PIR: JH0392; OKHURL.
DR HSSP: P00514; IAPK.
DR MIM: 176911; -.
DR PROSITE: PS00888; CNMP_BINDING_1.
DR PROSITE: PS00889; CNMP_BINDING_2.
KW CAMP-BINDING; PHOSPHORYLATION; DUPLICATION; MULTIGENE FAMILY;
KW ACETYLATION.
FT INIT_MET 0 0 BY SIMILARITY.
FT DOMAIN 1 135 DIMERIZATION AND PHOSPHORYLATION.
FT SITE 95 99 PSEUDOPHOSPHORYLATION MOTIF.
FT NP_BIND 136 253 CAMP.
FT NP_BIND 254 380 CAMP.
FT REPEAT 136 253
FT REPEAT 254 380
FT MOD_RES 1 1 ACETYLATION (PROBABLE).
FT DISULFID 17 17 INTERCHAIN (WITH 38 AND 17)
(BY SIMILARITY).
FT DISULFID 38 38 INTERCHAIN (WITH 38 AND 17)
(BY SIMILARITY).
FT MOD_RES 100 100 PHOSPHORYLATION (BY SIMILARITY).
FT BINDING 201 201 CAMP.
FT BINDING 210 210 CAMP.
FT BINDING 325 325 CAMP.
FT BINDING 334 334 CAMP.
SQ SEQUENCE 380 AA; 43027 MW; 705B1C9E CRC32;

Query Match 66.2%; Score 51; DB 5; Length 380;
Best Local Similarity 66.7%; Pred. No. 6.00e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 vldxcivhl 37
III I::II
QY 1 VLKRCLLHL 9

```

```

RESULT 9
ID CG2A_BOVIN STANDARD; PRT; 406 AA.
AC P30274;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE G2/MITOTIC-SPECIFIC CYCLIN A (FRAGMENT).
GN CCNA.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-LYMPHOCYTES;
RX MEDLINE: 93091274.
RA KOBAYASHI H., STEWART E., POON R., ADAMCZEWSKI J.P., GANNON J.,
RL MONT T.;
MOL. BIOL. CELL 3:1279-1294(1992).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G2/M
(MITOSIS) TRANSITION. INTERACTS WITH THE CDC2 AND CDK2 PROTEIN
KINASES TO FORM MPF. G2/M CYCLINS ACCUMULATE STEADILY DURING G2
AND ARE ABRUPTLY DESTROYED AT MITOSIS.
CC -1- SIMILARITY: TO OTHER CYCLINS. STRONGEST TO OTHER CYCLINS A.
EMBL: X68321; E54213; -.
DR PIR: S24788; S24788.
DR PROSITE: PS00292; CYCLINS.
KW CYCLIN; CELL CYCLE; CELL DIVISION; MITOSIS.
FT NON_TER 1 1
SQ SEQUENCE 406 AA; 45968 MW; 3CC9F046 CRC32;

Query Match 66.2%; Score 51; DB 2; Length 406;
Best Local Similarity 75.0%; Pred. No. 6.00e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 361 lkpcldl 368
III I::II
QY 2 LKRCLLHL 9

RESULT 10
ID E1BL_ADE41 STANDARD; PRT; 472 AA.
AC P10546;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE E1B PROTEIN, LARGE T-ANTIGEN (EARLY E1B 52 KD PROTEIN).
OS HUMAN ADENOVIRUS TYPE 41.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; ADENOVIRIDAE; MASTADENOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88084437.
RA VAN LOON A.E., LIGTENBERG M., REEMST A.M.C.B., SUSSENBACH J.S.,
RL ROZIJN T.H.;
GENE 58:109-126(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92230230.
RA ALLARD A., WADELL G.;
RL VIROLOGY 188:319-330(1992).
DR EMBL: M18289; G209781; -.
DR EMBL: M87544; G209892; -.
DR PIR: F27333; WMADF6.
KW EARLY PROTEIN.
FT CONFLICT 54 55 MISSING (IN REF. 2).
FT CONFLICT 172 173 TV -> IL (IN REF. 2).
FT CONFLICT 234 234 S -> P (IN REF. 2).
SQ SEQUENCE 472 AA; 52157 MW; 7B9EE276 CRC32;

Query Match 66.2%; Score 51; DB 3; Length 472;
Best Local Similarity 55.6%; Pred. No. 6.00e+00;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 367 mlmrcctmhl 375

```



QY 1 VLKRCLLHL 9  
:|:|:|:|:|

RESULT 11  
ID YHC3\_YEAST STANDARD; PRT; 1146 AA.  
AC P38742;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 130.0 KD PROTEIN IN SNF6-SPO11 INTERGENIC REGION.  
GN YH023C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE; 94378003.  
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,  
RA DU Z., FAVELLO A., FULTON L., GATTUNG S., GEISEL C., KIRSTEN J.,  
RA KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,  
RA LATEILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,  
RA NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASKIS E., VAUGHAN K.,  
RA VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,  
RA VAUDIN M.;  
RL SCIENCE 265:2077-2082(1994).  
DR EMBL; U11582; G508759; -.  
DR PIR; S46837; S46837.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 1146 AA; 129972 MW; 08B60453 CRC32;

Query Match 66.2%; Score 51; DB 11; Length 1146;  
Best Local Similarity 55.6%; Pred. No. 6.00e+00;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 820 l1rscalh 828  
:|:|:|:|:|

QY 1 VLKRCLLHL 9

RESULT 12  
ID YKBO\_YEAST STANDARD; PRT; 1483 AA.  
AC P33202;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 167.8 KD PROTEIN CCE1-CAP1 INTERGENIC REGION.  
GN YK010C OR YK1162.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93127732.  
RA PASCOLO S., GHAZVINI M., BOYER J., COLLEAUX L., THIERRY A., DUJON B.;  
RL YEAST 8:987-995(1992).  
CC -!- A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-THIOLESTER  
CC FORMATION.  
CC -!- SIMILARITY: CONTAINS AN HECT-TYPE E3 UBIQUITIN-PROTEIN LIGASE  
CC DOMAIN.  
CC -!- SIMILARITY: SOME, TO S.POMBE SPAC12B10.01C.  
DR EMBL; S53418; G263499; -.  
DR EMBL; 228010; G485993; -.  
DR PIR; S30015; S30015.  
KW HYPOTHETICAL PROTEIN; UBIQUITIN CONJUGATION; LIGASE.  
FT DOMAIN 898 901  
FT POLY-GLU.  
FT DOMAIN 1376 1483  
FT BINDING 1450 1450 UBIQUITIN (BY SIMILARITY).  
SQ SEQUENCE 1483 AA; 167841 MW; 1A524F93 CRC32;

Query Match 66.2%; Score 51; DB 11; Length 1483;  
Best Local Similarity 62.5%; Pred. No. 6.00e+00;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 743 vlkkcifi 750  
:|:|:|:|:|

QY 1 VLKRCLLHL 8

RESULT 13  
ID Y388\_MYCGE STANDARD; PRT; 101 AA.  
AC P47628;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN MG388.  
GN MG388.  
OS MYCOPLASMA GENITALIUM.  
OC PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;  
OC MYCOPLASMATACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE; 96026346.  
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,  
RA FLEISCHMANN R.D., BULT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,  
RA FRITZMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,  
RA NGUYEN D.T., UUTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,  
RA TOMB J.F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,  
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;  
RL SCIENCE 270:397-403(1995).  
DR EMBL; U39724; G1046100; -.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 101 AA; 11930 MW; 10F7DD9C CRC32;

Query Match 64.9%; Score 50; DB 10; Length 101;  
Best Local Similarity 62.5%; Pred. No. 9.27e+00;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 57 lltcrlih 64  
:|:|:|:|:|

QY 1 VLKRCLLHL 8

RESULT 14  
ID RPEC\_SYNPY STANDARD; PRT; 209 AA.  
AC Q02185;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)  
DE BILIN BIOSYNTHESIS PROTEIN RPEC.  
GN RPEC.  
OS SYNECHOCOCUS SP. (STRAIN WH8020).  
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;  
OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93144698.  
RA DE LORIMIER R., WILBANKS S.M., GLAZER A.N.;  
RL PLANT MOL. BIOL. 21:225-237(1993).  
CC -!- FUNCTION: AN ENZYME INVOLVED IN THE BIOSYNTHESIS OF BILIN. MIGHT  
CC BE INVOLVED IN THE SPECIFIC ATTACHMENT OF PHYCOERYTHROBLIN (PEB)  
CC TO THE R-PHYCOCYANIN II ALPHA CHAIN.  
CC -!- SIMILARITY: BELONGS TO THE CPCF/RPCF/PECF FAMILY.  
DR EMBL; M95288; G154569; -.  
DR PIR; S31071; S31071.  
KW PHYCOBILISOME; LYASE.  
SQ SEQUENCE 209 AA; 22956 MW; 288A2226 CRC32;

Query Match 64.9%; Score 50; DB 8; Length 209;  
Best Local Similarity 75.0%; Pred. No. 9.27e+00;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 137 lercallal 144  
:|:|:|:|:|

QY 2 LKRCLLHL 9

```

RESULT 15
ID MYCN_HUMAN STANDARD; PRT; 464 AA.
AC P04198;
DT 20-MAR-1987 (REL. 04, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE N-MYC PROTO-ONCOGENE PROTEIN.
GN MYCN OR MYC.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86149381.
RA STANTON L.W., SCHWAB M., BISHOP J.M.;
RL PROC. NATL. ACAD. SCI. U.S.A. 83:1772-1776(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86092232.
RA KOHL N.E., LEGOUY E., DEPINHO R.A., NISEN P.D., SMITH R.K.,
RA GEE C.E., ALT F.W.;
RL NATURE 319:73-77(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88202932.
RA IBSON J.M., RABBITTS P.H.;
RL ONCOGENE 2:399-402(1988).
RN [4]
RP SEQUENCE OF 71-464 FROM N.A.
RX MEDLINE; 86179901.
RA SLAMON D.J., BOONE T.C., SEEGER R.C., KEITH D.E., CHAZIN V., LEE H.C.,
RA SOUZA L.M.;
RL SCIENCE 232:768-772(1986).
RN [5]
RP SEQUENCE OF 313-464 FROM N.A.
RX MEDLINE; 85215633.
RA MICHITSCH R.W., MELERA P.W.;
RL NUCLEIC ACIDS RES. 13:2545-2558(1985).
RN [6]
RP PHOSPHORYLATION BY CK-II.
RX MEDLINE; 93049352.
RA HAGIWARA T., NAKAYA K., NAKAMURA Y., NAKAJIMA H., NISHIMURA S.,
RA TAYA Y.;
RL EUR. J. BIOCHEM. 209:945-950(1992).
CC -!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED DURING FETAL DEVELOPMENT.
CC -!- DISEASE: AMPLIFICATION OF THE N-MYC GENE IS ASSOCIATED WITH A
VARIETY OF HUMAN TUMORS, MOST FREQUENTLY NEUROBLASTOMA, WHERE THE
LEVEL OF AMPLIFICATION APPEARS TO INCREASE AS THE TUMOR
PROGRESSES.
CC -!- SIMILARITY: TO OTHER MEMBERS OF THE MYC FAMILY OF HELIX-LOOP-HELIX
TRANSCRIPTION FACTORS.
CC EMBL; X03294; G35079; -.
DR EMBL; X03295; G871358; -.
DR EMBL; M13241; G386983; ALT_INIT.
DR EMBL; M13228; G386982; -.
DR EMBL; X02363; -; NOT_ANNOTATED_CDS.
DR EMBL; Y00664; G35076; ALT_INIT.
DR PIR; A25744; TVHUM2.
DR PIR; A01355; TVHUMC.
DR PIR; A22937; A22937.
DR PIR; S02249; S02249.
DR MIM; 164840; -.
DR PROSITE; PS00038; HELIX LOOP HELIX.
KW NUCLEAR PROTEIN; DNA-BINDING; PROTO-ONCOGENE; PHOSPHORYLATION.
FT DOMAIN 262 278 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 381 434 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 433 454 LEUCINE-ZIPPER.
FT MOD_RES 261 261 PHOSPHORYLATION (BY CK-II).
FT MOD_RES 263 263 PHOSPHORYLATION (BY CK-II).
FT CONFLICT 227 227 A -> P (IN REF. 2).

```

```

ET CONFLICT 363 363 I -> V (IN REF. 3).
SQ SEQUENCE 464 AA; 49561 MW; EB8C3FC9 CRC32;

Query Match      64.9%; Score 50; DB 6; Length 464;
Best Local Similarity 83.3%; Pred No. 9.27e+00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 319 ilkrcl 324
   :|||||
QY 1 VLKRCL 6

```

Search completed: Tue Jun 10 11:25:25 1997  
Job time : 8 secs.

\*\*\*\*\*  
W A S E R L E H  
\*\*\*\*\* (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:27:35 1997; MasPar time 1.92 Seconds  
Tabular output not generated. 56.863 Million cell updates/sec

Title: >US-08-231-565A-35  
Description: (1-10) from US08231565A.pep  
Perfect Score: 74  
Sequence: 1 VLSPSPACQLV 10

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 16.093; Variance 49.915; scale 0.322

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	74	100.0	10	15	Melanoma-specific mut	8.90e-02
2	74	100.0	10	15	gp100 melanoma antige	8.90e-02
3	74	100.0	661	14	Melanoma associated a	8.90e-02
4	74	100.0	661	15	MART-1 melanoma antig	8.90e-02
5	74	100.0	661	15	MART-1 melanoma antig	8.90e-02
6	67	90.5	9	15	Melanoma-specific mut	5.72e-01
7	56	75.7	1277	10	Plasmid pASK60-Strep	9.77e+00
8	48	64.9	1728	3	Deleted in Colorectal	7.05e+01
9	45	60.8	1290	15	Phospholipase C-gamma	1.44e+02
10	45	60.8	1830	8	Urea amide lyase.	1.44e+02
11	44	59.5	190	4	Sheep Corticotropin r	1.83e+02
12	44	59.5	480	1	Soluble intercellular	1.83e+02
13	44	59.5	507	9	ICAM-1.	1.83e+02
14	44	59.5	531	15	ICAM-1 CD54.	1.83e+02
15	44	59.5	532	11	Inter-cellular adhesi	1.83e+02
16	44	59.5	532	14	ICAM-1.	1.83e+02
17	44	59.5	532	1	Inter-cellular adhesio	1.83e+02
18	44	59.5	532	1	Inter-cellular adhesi	1.83e+02
19	44	59.5	532	16	Intracellular adhesio	1.83e+02
20	44	59.5	532	9	Human ICAM-1.	1.83e+02

21	44	59.5	532	7	R35071	ICAM-1.	1.83e+02
22	44	59.5	532	18	R91437	Human ICAM-1.	1.83e+02
23	44	59.5	532	4	R20809	Intercellular Adhesio	1.83e+02
24	44	59.5	562	1	P80458	Sequence of human int	1.83e+02
25	44	59.5	680	9	R48037	TICAM(453)IgG immuno	1.83e+02
26	44	59.5	704	1	P80087	Sequence of 85 kd pro	1.83e+02
27	44	59.5	740	8	R39683	VCAM/ICAM-3.	1.83e+02
28	43	58.1	71	8	R38795	Conotoxin prepropepti	2.31e+02
29	43	58.1	370	18	R98127	Human lymphocyte cell	2.31e+02
30	43	58.1	371	18	R98114	Human lymphocyte cell	2.31e+02
31	43	58.1	371	18	R98112	Human lymphocyte cell	2.31e+02
32	43	58.1	371	18	R98113	Human lymphocyte cell	2.31e+02
33	43	58.1	371	18	R98111	Human lymphocyte cell	2.31e+02
34	43	58.1	371	18	R98110	Human lymphocyte cell	2.31e+02
35	43	58.1	371	18	R98109	Human lymphocyte cell	2.31e+02
36	43	58.1	371	18	R98116	Human lymphocyte cell	2.31e+02
37	43	58.1	371	18	R98117	Human lymphocyte cell	2.31e+02
38	43	58.1	371	18	R98118	Human lymphocyte cell	2.31e+02
39	43	58.1	371	18	R98119	Human lymphocyte cell	2.31e+02
40	43	58.1	371	18	R98121	Human lymphocyte cell	2.31e+02
41	43	58.1	372	18	R98106	Human lymphocyte cell	2.31e+02
42	43	58.1	372	14	R83050	Human LHR.	2.31e+02
43	43	58.1	385	6	R28753	Human LAM-1.	2.31e+02
44	43	58.1	385	6	R32707	LAM-1 from pLAM-1.	2.31e+02
45	43	58.1	858	2	P70099	Sequence of elongatio	2.31e+02

ALIGNMENTS

RESULT 1  
ID R82171 standard; peptide; 10 AA.  
AC R82171;  
DT 25-MAR-1996 (first entry)  
DE Melanoma-specific mutant immunogen epitope 10mer peptide.  
KW Melanoma; Immunogen; epitope; homologue; vaccine; immunotherapy;  
OS cytotoxic T cell; lymphocyte; HLA-A2.  
KW Homo sapiens.  
PN WO9522561-A2.  
PD 24-AUG-1995.  
PF 16-FEB-1995; U01991.  
PR 16-FEB-1994; US-197399.  
PR 29-APR-1994; US-234784.  
PA (UYVI-Q) UNIV VIRGINIA PATENT FOUND.  
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
DR WPI; 95-302688/39.  
PT Melanoma-specific immunogen comprises epitope(s) homologous with  
PT pMel17 - are highly potent stimulators of HLA-A2+CTL's useful in  
PT adoptive immuno-therapy  
PS Example 8; Page 52; 148pp; English.  
CC A melanoma-specific immunogen homologous with pMel-17 comprises one  
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
CC R82108 are of particular interest. The immunogen can be used for  
CC partial protection in mammals against melanoma peptides which are  
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+  
CC CTLs in several cell lines and can be used in immunotherapy or  
CC incorporated into immunogenic conjugates as vaccines.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 74; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 8.90e-02; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

Db 1 vlspspacqlv 10  
| | | | |  
QY 1 VLSPSPACQLV 10

RESULT 2  
ID R84203 standard; Peptide; 10 AA.  
AC R84203;  
DT 25-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G10-8).

KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 55; Page 131; 184pp; English.  
 CC The immunogenic peptide is derived from cDNA25 (R84854), a  
 CC melanoma antigen derivative of gp100 (see R84855). The  
 CC peptide and its derivatives (see R84200-R84211) are used in  
 CC medicaments (vaccines) for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against melanoma-specific  
 CC antigens and its immunogenic peptides may be used in the  
 CC detection and isolation of the antigen from a sample, the  
 CC detection of which is indicative of a disease state  
 CC (melanoma or metastatic melanoma).  
 SQ Sequence 10 AA;

Query Match 100.0%; Score 74; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 8.90e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vlspacqlv 10  
 |||||  
 QY 1 VLPSPACQLV 10

RESULT 3  
 ID R78646 standard; Protein; 661 AA.  
 AC R84855;  
 DT 22-JAN-1996 (first entry)  
 DE Melanoma associated antigen gp100.  
 KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 OS Homo sapiens.  
 PN EP-668350-A1.  
 PD 23-AUG-1995.  
 PF 14-FEB-1995; 200348.  
 PR 16-FEB-1994; EP-200337.  
 PR 21-DEC-1994; EP-203709.  
 PA (ALKU ) AKZO NOBEL NV.  
 PI Adema GJ, Figdor CG;  
 DR WPI; 95-284790/38.  
 DR N-PSDB; Q96055.  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PT detection of tumours  
 PS Claim 1; Page 22-24; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC may be used in the production of vaccines. Nucleotide sequences  
 CC encoding the immunogenic peptides may be used as primers and probes  
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
 CC capable of binding to the melanoma associated antigen can be  
 CC cultured ex vivo and returned to melanoma particles, and when  
 CC radiolabelled, they may be used to identify tumour deposits.  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 74; DB 14; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 8.90e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 544 vlspacqlv 553  
 |||||  
 QY 1 VLPSPACQLV 10

RESULT 4  
 ID R84855 standard; Protein; 661 AA.  
 AC R84855;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen gp100.  
 KW gp100; MART-1; melanoma antigen recognised by T-cell;  
 KW cDNA25 antigen derivative; melanocyte; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian sp.

Key Location/Qualifiers  
 FT Peptide 154..163  
 FT /label= G9-154\_immunogenic\_peptide  
 FT Peptide 208..217  
 FT /label= G9-209\_immunogenic\_peptide  
 FT /note= "see R84210"  
 FT Peptide 280..288  
 FT /label= G9-280\_immunogenic\_peptide  
 FT /note= "see R84208"  
 FT Peptide 457..266  
 FT /label= immunogenic\_peptide  
 FT Peptide 476..485  
 FT /label= immunogenic\_peptide  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 7A; 184pp; English.  
 CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which  
 CC is recognized by T-lymphocytes, and is a derivative of the  
 CC melanoma-specific antigen cDNA25 (see R84854). gp100 is a source  
 CC of immunogenic peptides which are optionally modified to enhance  
 CC their binding to a MHC molecule, and used in medicaments,  
 CC especially vaccines, for the treatment or prevention (by  
 CC immunisation) of melanoma. Antibodies against cDNA25 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 74; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 8.90e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 544 vlspacqlv 553  
 |||||  
 QY 1 VLPSPACQLV 10

RESULT 5  
 ID R84854 standard; Protein; 661 AA.  
 AC R84854;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen cDNA25.  
 KW cDNA25; MART-1; melanoma antigen recognised by T-cell;  
 KW gp100 antigen derivative; melanoma; metastatic melanoma;  
 KW tumour-associated antigen; immunogen; diagnosis; prognosis;  
 KW prophylaxis; therapy; vaccine.  
 OS Mammalian sp.  
 FT Key  
 FT Location/Qualifiers  
 FT Peptide 457..466  
 FT /label= antigenic\_peptide

FT /note= "see R84199"  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI: 95-382963/49.  
 DR N-PSDB; T02716.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 5A; 184pp; English.  
 CC cDNA2 is a melanoma antigen (MART-1) which is recognized by  
 CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-  
 CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source  
 CC of immunogenic peptides (see R84199) which are optionally modified  
 CC (see R84200-R84211) to enhance their binding to a MHC molecule and  
 CC used in medicaments, especially vaccines, for the treatment or  
 CC prevention (by immunisation) of melanoma. Antibodies against cDNA2  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of the antigen from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 74; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 8.90e-02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 544 vlpspacqlv 553  
 QY 1 VLPSPACQLV 10  
 |||||

RESULT 6  
 ID R82128 standard; peptide; 9 AA.  
 AC R82128;  
 DT 25-MAR-1996 (first entry)  
 DE Melanoma-specific mutant immunogen epitope 9mer peptide.  
 KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
 KW cytotoxic T cell; lymphocyte; HLA-A2.  
 OS Homo sapiens.  
 PN WO9522561-A2.  
 PD 24-AUG-1995.  
 PF 16-FEB-1995; U01991.  
 PR 16-FEB-1994; US-197399.  
 PR 29-APR-1994; US-234784.  
 PA (UYVI-) UNIV VIRGINIA PATENT FOUND.  
 PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
 DR WPI: 95-302689/39.  
 PT Melanoma-specific immunogen comprises epitope(s) homologous with  
 PT pMel.17 - are highly potent stimulators of HLA-A2+CTL's useful in  
 PT adoptive immuno-therapy  
 PS Example 8; Page 51; 148pp; English.  
 CC A melanoma-specific immunogen homologous with pMel-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
 CC R82108 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pMel-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.  
 SQ Sequence 9 AA;

Query Match 90.5%; Score 67; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 5.72e-01;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vlpspacql 9  
 QY 1 VLPSPACQL 9  
 |||||

## RESULT 7

ID R52702 standard; Protein; 1277 AA.  
 AC R52702;  
 DT 11-JAN-1995 (first entry)  
 DE Plasmid pASK60-Strep reading frame "c" translation.  
 DE Streptavidin binding peptide; fusion protein; pASK60-Strep;  
 KW affinity chromatography; purification; peptide tag; detection.  
 OS Synthetic.  
 FH Key  
 FT Region 1..1277  
 FT /note= "translated from reading frame 'c' of  
 FT pASK60-Strep; all X's correspond to nonsense  
 FT codons"  
 FT Peptide 482..491  
 FT /label= streptavidin-binding peptide  
 PN GB2272698-A.  
 PD 25-MAY-1994.  
 PF 01-NOV-1993; 022501.  
 PR 03-NOV-1992; DB-237113.  
 PA (KUEH-) KUEHN KONSTRUKTION GMBH & CO KG KLAUS.  
 PA (BIOA-) INST BIOANALYTIK GEMEINNUTZIGE GMBH.  
 PI Schmidt T, Skerra A;  
 DR WPI: 94-153484/19.  
 DR N-PSDB; Q62676.  
 PT New fusion peptide(s) - have easily controlled binding properties  
 PT and are capable of binding to streptavidin  
 PS Example 5; Fig 8; 53pp; English.  
 CC Plasmid pASK60-Strep was produced starting from pASK40  
 CC (Biotechnology 9, 273-278, 1991) using site-directed mutagenesis  
 CC and PCR. The plasmid contains an improved set of unique restriction  
 CC sites, including two sites located directly at the 3' end of the  
 CC region coding for the OmpA signal peptide. The polylinker is  
 CC followed by a DNA sequence coding for a streptavidin binding  
 CC peptide. The plasmid pASK60-Strep is useful for the expression of  
 CC polypeptides fused to a streptavidin binding peptide. The  
 CC production of such fusion proteins can be specifically detected  
 CC using a streptavidin-alkaline phosphatase conjugate.  
 SQ Sequence 1277 AA;

Query Match 75.7%; Score 56; DB 10; Length 1277;  
 Best Local Similarity 77.8%; Pred. No. 9.77e+00;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 284 lpspgcalv 292  
 QY 2 LPSPCQLV 10  
 |||||

## RESULT 8

ID R13144 standard; Protein; 1728 AA.  
 AC R13144;  
 DT 04-OCT-1991 (first entry)  
 DE Deleted in Colorectal Carcinomas.  
 KW DCC gene; cancer; diagnosis; antibodies; tumorigenesis; neoplasm.  
 OS Homo sapiens.  
 FH Key  
 FT Protein 202..1648  
 FT /label= DCC  
 FT Peptide 202..227  
 FT /label= sig\_peptide  
 FT Protein 228..1648  
 FT /label= mat\_protein  
 PN WO9109964-A.  
 PD 11-JUL-1991.  
 PF 19-DEC-1990; U07314.  
 PR 04-JAN-1990; US-460981.  
 PA (UYJO ) JOHNS HOPKINS UNIV.  
 PI Vogelstein B.  
 DR WPI: 91-222913/30.  
 DR N-PSDB; Q12752.  
 PT Human DCC gene, deleted in colorectal carcinoma - and diagnosis  
 PT or prognosis of neoplasms by detecting loss of gene function or

PT expression prods., or mutation(s)  
 PS Claim 44; Page 31; 51pp; English.  
 CC Cells transformed with the wild-type DCC gene can be used as model  
 CC systems to study cancer remission and drug therapy. DCC polypeptide  
 CC expression prods. may be used to reverse the neoplastic state.  
 CC X1615 represent an amino acid illegible in the specification, all  
 CC other xs are encoded by stop codons.  
 CC See also Q12752-55.  
 SQ Sequence 1728 AA;

Query Match 64.9%; Score 48; DB 3; Length 1728;  
 Best Local Similarity 71.4%; Pred. No. 7.05e+01;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1465 ilpsptc 1471  
 :||||:  
 QY 1 VLSPAC 7

RESULT 9  
 ID R90583 standard; Protein; 1290 AA.  
 AC R90583;  
 DT 03-APR-1996 (first entry)  
 DE Phospholipase C-gamma-1.  
 KW Phospholipase C-gamma-1; PLC-gamma-1; phosphoinositide.  
 OS Rattus sp.  
 PN US5474921-A.  
 PD 12-DEC-1995.  
 PF 15-OCT-1993; 138641.  
 PR 15-OCT-1993; US-138641.  
 PA (MERI ) MERCK & CO INC.  
 PI Koblan KS, Pompilano DL;  
 DR WPI: 96-048545/05.  
 DR N-PSDB; T12292.  
 PT Method for expression and isolation of mammalian phospholipase  
 PT C-gamma-1 - useful for determining inhibitory activity of test  
 PT compounds towards phospho:inositide-specific phospholipase-C enzyme.  
 PS Claim 1; Column 13-20; 25pp; English.  
 CC Rat phosphoinositide-specific phospholipase C-gamma-1 (EC-3.1.4.3)  
 CC (R90583) is obtd. by expression in a transformed bacterial host of  
 CC cDNA (T12292) encoding rat PLC-gamma-1 and DNA coding for an epitope  
 CC tag (Glu-Glu-Phe) which is incorporated at the C-terminus of the  
 CC recombinant PLC-gamma-1 to facilitate affinity purification. The  
 CC recombinant PLC-gamma-1 is used to assay the inhibitory activity of  
 CC a test cpds. against PLC-gamma-1.  
 SQ Sequence 1290 AA;

Query Match 60.8%; Score 45; DB 15; Length 1290;  
 Best Local Similarity 66.7%; Pred. No. 1.44e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 875 vldvpacq1 883  
 :||||:  
 QY 1 VLSPACQ 9

RESULT 10  
 ID R44504 standard; Protein; 1830 AA.  
 AC R44504;  
 DT 17-JUN-1994 (first entry)  
 DE Urea amide lyase.  
 KW Urea amide lyase; URL; yeast; Candida; recombinant; plasmid;  
 KW transformation; urea determination.  
 OS Candida utilis.  
 PN J05304965-A.  
 PD 19-NOV-1993.  
 PF 27-APR-1992; 107799.  
 PR (TOYM ) TOYOBO KK.  
 PA WPI: 93-408324/51.  
 DR N-PSDB; Q53454.  
 PT DNA having genetic information of urea amide lyase - prepd. by  
 PT culturing Candida utilis

PS Claim 1; Page 10-17; 17pp; Japanese.  
 CC This sequence represents a protein having urea amide lyase (URL)  
 CC activity and is encoded by a DNA sequence which was isolated from a  
 CC yeast belonging to the Candida genus. The DNA encoding this sequence  
 CC maybe used in a recombinant plasmid to transform a microorganism, and  
 CC URL may be isolated from the culture medium. This method allows  
 CC production of URL of high purity. The URL prepared gives a high  
 CC sensitivity in urea determination.  
 SQ Sequence 1830 AA;

Query Match 60.8%; Score 45; DB 8; Length 1830;  
 Best Local Similarity 55.6%; Pred. No. 1.44e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 372 vlaytpcpl 380  
 :||||:  
 QY 1 VLSPACQ 9

RESULT 11  
 ID P40294 standard; Protein; 190 AA.  
 AC P40294;  
 DT 23-JUL-1992 (first entry)  
 DE Sheep Corticotropin releasing factor.  
 KW CRF; sheep hypothalamus.  
 OS Ovis aries.  
 PH Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /label= signal 144..147  
 FT Region  
 FT /note= "run of basic residues"  
 FT Region 148..188  
 FT /note= "CRF"  
 PN J59140884-A.  
 PD 13-AUG-1984.  
 PF 29-JAN-1983; 013511.  
 PR 29-JAN-1983; JP-013511.  
 PA (DAIN ) DAINIPPON PHARM KK.  
 PA (MITU ) MITSUBISHI CHEM. IND. KK.  
 DR WPI: 84-234742/38.  
 DR N-PSDB; N40238.  
 DR Cloning DNA fragment - from gene coding corticotropin-release factor  
 PS Disclosure; Fig 2; 15pp; Japanese.  
 CC Total RNA was extracted from sheep hypothalamus. The mRNA fraction  
 CC was reverse transcribed to cDNA. A cDNA library was constructed and  
 CC screened with a probe based on amino acids 167-171 of the known CRF  
 CC sequence. A cDNA clone which hybridised with the probe was itself  
 CC used to rescreen the library. A clone containing the entire coding  
 CC region for CRF was identified. See also N40237 and N40451.  
 SQ Sequence 190 AA;

Query Match 59.5%; Score 44; DB 4; Length 190;  
 Best Local Similarity 62.5%; Pred. No. 1.83e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 16 llpsppcr 23  
 :||||:  
 QY 1 VLSPACQ 8

RESULT 12  
 ID R06240 standard; protein; 480 AA.  
 AC R06240;  
 DT 07-DEC-1990 (first entry)  
 DE Soluble intercellular adhesion molecule-1 (sICAM-1).  
 KW Soluble intercellular adhesion molecule-1; Hela; LFA-1; rhinovirus;  
 KW Cxsackie A virus; ds;  
 OS Homo sapiens.  
 PN EP-379904-A.  
 PD 01-AUG-1990.  
 PF 12-JAN-1990; 100557.  
 PR 24-JAN-1989; US-301192.  
 PR 13-DEC-1989; US-445951.

PA (MOLE-) MOLECULAR THERAPEUT.  
 PI Greve J, McClelland A;  
 DR WPI; 90-232770/31.  
 DR N-PSDB; Q05536.  
 PT Human soluble inter-cellular adhesion molecule-1 - which inhibits  
 PT lymphocyte function associated with LFA-1 and ICAM-1 interaction  
 PS Claim 2; Fig 1; 16pp; English.  
 CC SICAM can inhibit lymphocyte function associated with antigen  
 CC LFA-1 and ICAM-1 interaction, useful in treating inflammation,  
 CC graft rejection, and for LFA-1 expressing tumour cells. Can also  
 CC inhibit infection by rhinovirus and Cocksackie A virus, increase  
 CC cell adhesion and reverse immune dysfunction due to excess sICAM-1.  
 SQ Sequence 480 AA;

Query Match 59.5%; Score 44; DB 1; Length 480;  
 Best Local Similarity 60.0%; Pred. No. 1.83e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 213 vlpapppqly 222  
 |||: : |||  
 QY 1 VLPSPACQLV 10

## RESULT 13

ID R48038 standard; Protein; 507 AA.  
 AC R48038;  
 DT 12-JUN-1994 (first entry)  
 DE ICAM-1.  
 KW ICAM-1; intercellular adhesion molecule-1; human rhinovirus; HRV;  
 KW Probe; primer; polymerase chain reaction; PCR; hybridation.  
 OS Homo sapiens.  
 PN WO9400485-A.  
 PD 06-JAN-1994.  
 PF 22-JUN-1993; U05972.  
 PR 22-JUN-1992; US-903069.  
 PA (MILE ) MILES INC.  
 PI Greve JM, McClelland A;  
 DR WPI; 94-026146/03.  
 PT Multimeric forms of inter-cellular adhesion mol. (ICAM) -  
 PT displaying enhanced binding of human rhinovirus and able to  
 PT reduce its infectivity  
 PS Fig 1; Page 37; 70pp; English  
 CC Oligonucleotides Q55327-35 were used to create forms of  
 CC ICAM-1 (R48038) that facilitate cross-linking and  
 CC multimerization. The primers given in Q55336-39 were used to clone  
 CC ICAM(185)/I96 immunoadhesin fusion protein. ICA(453)/I96 fusion  
 CC (R48037) in encoded by sequence Q55340. Mutein Q55342, a mutated  
 CC form of Q55340, encodes tICAM(452) cysteine mutant terminated at  
 CC residue 452. Probe Q55341 was used for HRV identification.  
 SQ Sequence 507 AA;

Query Match 59.5%; Score 44; DB 9; Length 507;  
 Best Local Similarity 60.0%; Pred. No. 1.83e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 188 vlpapppqly 197  
 |||: : |||  
 QY 1 VLPSPACQLV 10

## RESULT 14

ID R80110 standard; Protein; 531 AA.  
 AC R80110;  
 DT 04-MAR-1996 (first entry)  
 DE ICAM-1 CD54.  
 KW CD54; intercellular adhesion molecule-1; ICAM-1; integrin.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..25  
 FT /label= sig\_peptide  
 FT Region 41..100  
 FT /label= Ig1  
 FT Region 128..190

FT /label= Ig2 230..294  
 FT Region  
 FT /label= Ig3 325..375  
 FT Region  
 FT /label= Ig4 413..461  
 FT Region  
 FT /label= Ig5 481..503  
 FT Domain  
 FT /label= Trans  
 FT /note= "transmembrane domain"  
 FT Binding\_site 152..154  
 FT /note= "cell attachment site"  
 PN WO9528170-A1.  
 PD 26-OCT-1995.  
 PF 19-APR-1995; U04886.  
 PR 19-APR-1994; US-229513.  
 PA (UNIV ) UNIV KANSAS  
 PI Benedict S, Chan MA, Siahaan TJ, Tibbetts SA;  
 DR WPI; 95-373631/48.  
 PT Changing peptide reactivity via conjugation with a second peptide -  
 PT causes change in conformation of first peptide, pref. LFA-1 and  
 PT ICAM-1 functional domain derived peptide(s)  
 PS Claim 14; Page 38-40; 64pp; English.  
 CC Functional domains derived from LFA-1 alpha subunit CD11a (R80107),  
 CC LFA-1 beta subunit CD18 (R80107) and ICAM-1 CD43 (R80110) are used  
 CC as the basis of short, synthetic peptides (R80109, R80111-24) that  
 CC modulate ICAM/LFA binding interaction. CD18 is obt'd. from human  
 CC T-lymphocyte HL-60 cells.  
 SQ Sequence 531 AA;

Query Match 59.5%; Score 44; DB 15; Length 531;  
 Best Local Similarity 60.0%; Pred. No. 1.83e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 212 vlpapppqly 221  
 |||: : |||  
 QY 1 VLPSPACQLV 10

## RESULT 15

ID R58779 standard; Protein; 532 AA.  
 AC R58779;  
 DT 24-APR-1995 (first entry)  
 DE Inter-cellular adhesion molecule protein.  
 KW adhesion; transmembrane; anticoagulation; heparin; heparinoid;  
 KW proteoglycan; prostasin; inflammatory; immunological; lymphocytes;  
 KW multiple sclerosis; biliary; cirrhosis; erythematosis; rheumatism;  
 KW infarct; carcinoma.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..25  
 FT /note= "signal peptide"  
 FT Protein 26..533  
 FT /note= "mature protein"  
 FT Modified\_site 130..132  
 FT /note= "N-linked glycosylation site"  
 FT Domain 152..154  
 FT /note= "RGD-like domain"  
 FT Modified\_site 183..185  
 FT /note= "N-linked glycosylation site"  
 FT Modified\_site 202..204  
 FT /note= "N-linked glycosylation site"  
 FT Modified\_site 267..269  
 FT /note= "N-linked glycosylation site"  
 FT Modified\_site 296..298  
 FT /note= "N-linked glycosylation site"  
 FT Modified\_site 385..387  
 FT /note= "N-linked glycosylation site"  
 FT Modified\_site 406..408  
 FT /note= "N-linked glycosylation site"  
 FT Domain 481..503  
 FT /note= "Transmembrane domain"  
 PN WO9418988-A.

PD 01-SEP-1994.  
 PF 22-FEB-1994; E00506.  
 PR 22-FEB-1993; EP-102750.  
 PA (VARN/) VON ARNIM U.  
 PI Von ARNIM U.  
 DR WPI; 94-293982/36.  
 DR N-PSDB; Q67439.  
 PT Pharmaceutical for treatment or inflammatory or immunological  
 PT diseases - contg. heparins, heparinoids, proteoglycans and/or  
 PT prostavasin  
 PS Disclosure; Fig 6a; 37pp; English.  
 CC The protein contains a signal peptide sequence (a.a. 1-25). The  
 CC mature protein also contains a transmembrane domain (a.a. 481-503).  
 CC There is an RGD-like sequence (a.a. 152-154). The protein is N-linked  
 CC glycosylated at 7 sites (a.a. 130-132, 183-185, 202-204, 267-269,  
 CC 296-298, 385-387, 406-408). Based on known immunoglobulin related  
 CC sequences, the ICAM protein can be divided into 5 domains (a.a. 53-137,  
 CC 140-231, 337-416 and 424-502). These domains have significant similarity  
 CC to members of the immunoglobulins. The mature protein is a transmembrane  
 CC protein which binds activated T lymphocytes. Diseases of the system are  
 CC inflammatory and/or immunological including multiple sclerosis, lupus  
 CC erythematosus, primary biliary cirrhosis and other systemic carcinoma  
 CC diseases. Multiple sclerosis patients have been shown to encode a  
 CC hypervariable ICAM. The binding of lymphocytes to cells expressing ICAM  
 CC can be used to search anticoagulant factors to treat the above mentioned  
 CC diseases.  
 SQ Sequence 532 AA;

Query Match 59.5%; Score 44; DB 11; Length 532;  
 Best Local Similarity 60.0%; Pred. No. 1.83e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 213 vlpattppqlv 222  
 |||: |||  
 QY 1 VLFSPACQLV 10

Search completed: Tue Jun 10 11:27:44 1997  
 Job time : 9 secs.



\*\*\*\*\*

WIREH

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:27:06 1997; MasPar time 2.83 Seconds  
Tabular output not generated. 100.872 Million cell updates/sec

Title: >US-08-231-565A-35  
Description: (1-10) from US08231565A.pep  
Perfect Score: 74  
Sequence: 1 VLPSPACQLV 10

Scoring table: PAM 150  
Gap 15

Searched: 89912 seqs, 28507787 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir50

1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 22.099; Variance 31.490; scale 0.702

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	74	100.0	661 13	glycoprotein gp100 P	7.80e-05
2	74	100.0	668 13	melanocyte-specific	7.80e-05
3	62	83.8	491 14	melanoma antigen hom	2.92e-02
4	58	78.4	626 14	Pmel 17 protein - mo	1.90e-01
5	57	77.0	1042 13	mucin (clone L31) -	3.00e-01
6	57	77.0	1056 13	tracheobronchial muc	3.00e-01
7	50	67.6	2704 12	G surface protein 16	6.56e+00
8	50	67.6	2718 12	G surface protein -	6.56e+00
9	50	67.6	3020 13	mucin 2 precursor, i	6.56e+00
10	49	66.2	392 1	naringenin-chalcone	1.00e+01
11	49	66.2	568 12	ubiquinol--cytochrom	1.00e+01
12	49	66.2	1015 13	Human giant larvae h	1.00e+01
13	48	64.9	148 9	fur protein - Yersin	1.52e+01
14	48	64.9	671 10	hypothetical protein	1.52e+01
15	48	64.9	1427 13	tumor suppressor - A	1.52e+01
16	48	64.9	1447 13	tumor suppressor pro	1.52e+01
17	47	63.5	312 14	Cdk-activating kinase	2.29e+01
18	47	63.5	346 14	MO15/CDK-activating	2.29e+01
19	47	63.5	346 14	cell division cycle-	2.29e+01
20	47	63.5	346 13	CDK-activating prote	2.29e+01
21	47	63.5	407 14	[pyruvate dehydrogen	2.29e+01

RESULT ENTRY	22	47	63.5	407 13	170159	pyruvate dehydrogena	2.29e+01
TITLE	23	47	63.5	612 14	S23174	endothelial leukocyt	2.29e+01
ALTERNATE_NAMES	24	47	63.5	1745 12	B42755	E-selectin precursor	2.29e+01
ORGANISM	25	47	63.5	20 8	S44816	F44E2.1 protein - Ca	2.29e+01
DATE	26	46	62.2	20 8	B44920	2-halobenzoate 1,2-d	3.42e+01
ACCESSIONS	27	46	62.2	121 8	S28738	hypothetical protein	3.42e+01
REFERENCE	28	46	62.2	216 11	S28100	hypothetical protein	3.42e+01
#authors	29	46	62.2	223 14	JX0222	ubiquitin carboxyl-t	3.42e+01
#journal	30	46	62.2	291 3	EEWTG	gamma-gliadin B prec	3.42e+01
#title	31	46	62.2	308 9	S40873	hypothetical protein	3.42e+01
#accession	32	46	62.2	323 8	JC4774	fiber protein - huma	3.42e+01
#molecule_type	33	46	62.2	360 16	S38570	glyceraldehyde-3-pho	3.42e+01
#residues	34	46	62.2	391 10	E64474	hypothetical protein	3.42e+01
#authors	35	46	62.2	1134 14	A41350	adenylate cyclase (E	3.42e+01
#journal	36	45	60.8	72 13	I68511	small proline-rich p	5.10e+01
#title	37	45	60.8	92 4	MNVUSH	nonstructural protei	5.10e+01
#accession	38	45	60.8	305 8	S55655	ribonucleotide reduc	5.10e+01
REFERENCE	39	45	60.8	475 8	A47674	alpha-amylase (EC 3.	5.10e+01
#authors	40	45	60.8	611 14	A53418	calmeglin precursor -	5.10e+01
#journal	41	45	60.8	837 14	A42112	mucin-like peptide M	5.10e+01
#title	42	45	60.8	1290 13	A36466	1-phosphatidylinosit	5.10e+01
#accession	43	45	60.8	1290 14	A31317	1-phosphatidylinosit	5.10e+01
REFERENCE	44	45	60.8	1291 14	S00666	1-phosphatidylinosit	5.10e+01
#authors	45	45	60.8	3343 12	S44887	ZK112.7 protein - Ca	5.10e+01

## ALIGNMENTS

RESULT 1

ENTRY

TITLE

ALTERNATE\_NAMES

ORGANISM

DATE

ACCESSIONS

REFERENCE

#authors

#journal

#title

#accession

#molecule\_type

#residues

#authors

#journal

#title

#accession

#status

#molecule\_type

#residues

KEYWORDS

SUMMARY

Query Match

Best Local Similarity

Matches

Db

QY

RESULT 2

ENTRY

TITLE

```

ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS    30-Sep-1993
REFERENCE      A41234
#authors      Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
              Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
              Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
              coat color locus on mouse chromosome 10 and is in a
              syntenic region on human chromosome 12.
#cross-references MUID:92021023
#accession    A41234
##status      preliminary
##molecule_type mRNA
##residues    1-668 #label KWO
##cross-references GB:M77348
SUMMARY       #length 668 #molecular-weight 70932 #checksum 6409
Query Match   100.0%; Score 74; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 7.80e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 544 vlpspacqlv 553
|||||
QY 1 VLPSACQLV 10

RESULT 3
ENTRY      A49179          #type fragment
TITLE      melanoma antigen homolog rpe1 - bovine (fragment)
ORGANISM   #formal_name Bos primigenius taurus #common_name cattle
DATE       19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
23-Mar-1995
ACCESSIONS A49179
REFERENCE   A49179
#authors    Kim, R.Y.; Wistow, G.J.
#journal     Exp. Eye Res. (1992) 55:657-662
#title       The cDNA RPE1 and monoclonal antibody HMB-50 define gene
              products preferentially expressed in retinal pigment
              epithelium.
#cross-references MUID:93122163
#accession  A49179
##status    preliminary
##molecule_type nucleic acid
##residues  1-491 #label KIM
##cross-references NCBIN:122438; NCBIP:122439
##experimental_source retinal pigment epithelium
##note       #length 491 #checksum 3125
SUMMARY     #length 491 #checksum 3125
Query Match   83.8%; Score 62; DB 14; Length 491;
Best Local Similarity 90.0%; Pred. No. 2.92e-02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 372 vlpspacqlv 381
|||||
QY 1 VLPSACQLV 10

RESULT 4
ENTRY      S53871          #type complete
TITLE      Pmel 17 protein - mouse
ORGANISM   #formal_name Mus musculus #common_name house mouse
DATE       27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
03-Nov-1995
ACCESSIONS S53871
REFERENCE   S53871
#authors    Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
              Chintamaneni, C.; Bennett, D.; Pickard, R.T.
#journal     Nucleic Acids Res. (1995) 23:154-158
#title       Mouse silver mutation is caused by a single base insertion in

```

```

#accession    S53871          the putative cytoplasmic domain of Pmel 17.
##status      preliminary
##molecule_type mRNA
##residues    1-626 #label KWO
SUMMARY       #length 626 #molecular-weight 65979 #checksum 5710
Query Match   78.4%; Score 58; DB 14; Length 626;
Best Local Similarity 80.0%; Pred. No. 1.90e-01;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 511 vlpspacqlv 520
|||||
QY 1 VLSPACQLV 10

RESULT 5
ENTRY      A57534          #type fragment
TITLE      mucin (clone I31) - human (fragment)
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change
06-Sep-1996
ACCESSIONS A57534
REFERENCE   A57534
#authors    Lesuffleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.;
              Swallow, D.M.; Zweibaum, A.; Real, F.X.
#journal     J. Biol. Chem. (1995) 270:13665-13673
#title       Characterization of a mucin cDNA clone isolated from HT-29
              mucus-secreting cells. The 3' end of MUC5AC?
#accession  A57534
##status      preliminary
##molecule_type mRNA
##residues    1-1042 #label LES
##cross-references GB:Z48314
GENETICS     #gene          GDB:MUC5AC
              ##cross-references GDB:454136
              #map_position 11p15.5-11p15.5
SUMMARY       #length 1042 #checksum 4798
Query Match   77.0%; Score 57; DB 13; Length 1042;
Best Local Similarity 66.7%; Pred. No. 3.00e-01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 438 lpspicqli 446
|||||
QY 2 LPSPACQLV 10

RESULT 6
ENTRY      A53767          #type complete
TITLE      tracheobronchial mucin MUC5 - human
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
06-Sep-1996
ACCESSIONS A53767
REFERENCE   A53767
#authors    Meerzaman, D.; Charles, P.; Daskal, E.; Polymeropoulos, M.H.;
              Martin, B.M.; Rose, M.C.
#journal     J. Biol. Chem. (1994) 269:12932-12939
#title       Cloning and analysis of cDNA encoding a major airway
              glycoprotein, human tracheobronchial mucin (MUC5).
#accession  A53767
##status      preliminary
##molecule_type mRNA
##residues    1-1056 #label MEE
##cross-references GB:U06711
##note        authors translated the codon TAT for residue 1054 as Thr
GENETICS     #gene          GDB:MUC5B; MUC5
              ##cross-references GDB:128436
              #map_position 11p15.5-11p15.5
SUMMARY       #length 1056 #molecular-weight 113042 #checksum 6706

```

```

Query Match          77.0%; Score 57; DB 13; Length 1056;
Best Local Similarity 66.7%; Pred. No. 3.00e-01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 528 lpspachli 536
|||||l:l:
2 LPSAPACLV 10

RESULT 7
ENTRY S09118 #type complete
TITLE G surface protein 168 - Paramacium primaurelia (SGC5)
ORGANISM #formal_name Paramacium primaurelia
DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
31-Dec-1993
ACCESSIONS S09118
REFERENCE S09118
#authors Prat, A.
#journal J. Mol. Biol. (1990) 211:521-535
#title Conserved sequences flank variable tandem repeats in two
alleles of the G surface protein of Paramacium primaurelia.
#cross-references MUID:90172419
#accession S09118
#molecule_type DNA
#residues 1-2704 #label PRA
#cross-references EMBL:X52133
GENETICS
#genetic_code SGC5 #length 2704 #molecular-weight 278773 #checksum 3244
SUMMARY

Query Match          67.6%; Score 50; DB 12; Length 2704;
Best Local Similarity 85.7%; Pred. No. 6.56e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1876 lpapacq 1882
|||||
2 LPSAPACQ 8

RESULT 8
ENTRY A23475 #type complete
TITLE G surface protein - Paramacium primaurelia (SGC5)
ORGANISM #formal_name Paramacium primaurelia
DATE 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change
31-Dec-1993
ACCESSIONS A23475
REFERENCE A23475
#authors Prat, A.; Katinka, M.; Caron, F.; Meyer, E.
#journal J. Mol. Biol. (1986) 189:47-60
#title Nucleotide sequence of the Paramacium primaurelia G surface
protein. A huge protein with a highly periodic structure.
#cross-references MUID:87060934
#accession A23475
#molecule_type DNA
#residues 1-2718 #label PRA
#note the authors translated the codon TGC for residue 2665 as
Trp
GENETICS
#genetic_code SGC5 #length 2718 #molecular-weight 279853 #checksum 5861
SUMMARY

Query Match          67.6%; Score 50; DB 12; Length 2718;
Best Local Similarity 85.7%; Pred. No. 6.56e+00;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1887 lpapacq 1893
|||||
2 LPSAPACQ 8

RESULT 9
ENTRY A43932 #type fragments

```

```

TITLE mucin 2 precursor, intestinal - human (fragments)
ALTERNATE_NAMES mucin SMOG-41
ORGANISM #formal_name Homo sapiens #common_name man
DATE 10-Mar-1993 #sequence_revision 12-Apr-1996 #text_change
06-Sep-1996
ACCESSIONS A49963; A45106; B45106; A43932; B33532; A61257; PQ0328;
PQ0329
REFERENCE A49963
#authors Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Siddiki, B.; Kim,
Y.S.
#journal J. Biol. Chem. (1994) 269:2440-2446
#title Molecular cloning of human intestinal mucin (MUC2) cDNA.
Identification of the amino terminus and overall sequence
similarity to prepro-von Willebrand factor.
#accession A49963
#molecule_type mRNA
#residues 1-639 #label GU1
#cross-references GB:L21998
REFERENCE A45106
#authors Gum Jr., J.R.; Hicks, J.W.; Toribara, N.W.; Rothe, E.M.;
Lagace, R.E.; Kim, Y.S.
#journal J. Biol. Chem. (1992) 267:21375-21383
#title The human MUC2 intestinal mucin has cysteine-rich subdomains
located both upstream and downstream of its central
repetitive region.
#cross-references MUID:93016075
#accession A45106
#status not compared with conceptual translation
#molecule_type mRNA
#residues 626-1895 #label GU2
#cross-references NCBI:P116706
#note sequence extracted from NCBI backbone
#accession B45106
#status not compared with conceptual translation
#molecule_type mRNA
#residues 2037-3020 #label GU3
#cross-references NCBI:P116698
#experimental_source colon
#note sequence extracted from NCBI backbone
REFERENCE A43932
#authors Toribara, N.W.; Gum Jr., J.R.; Culhane, P.J.; Lagace, R.E.;
Hicks, J.W.; Petersen, G.M.; Kim, Y.S.
#journal J. Clin. Invest. (1991) 88:1005-1013
#title MUC-2 human small intestinal mucin gene structure. Repeated
arrays and polymorphism.
#cross-references MUID:91358717
#accession A43932
#molecule_type DNA
#residues 1343-1350,'L',1352-1411,'S',1413-1448,'P',1450-1503,'T',
1505-1915 #label TOR
#cross-references NCBI:N55749; NCBI:P:55750
#note sequence inconsistent with the nucleotide translation
#note sequence extracted from NCBI backbone
REFERENCE A33532
#authors Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampport,
D.T.A.; Kim, Y.S.
#journal J. Biol. Chem. (1989) 264:6480-6487
#title Molecular cloning of human intestinal mucin cDNAs. Sequence
analysis and evidence for genetic polymorphism.
#cross-references MUID:89197956
#accession B33532
#molecule_type mRNA
#residues 1916-2193 #label GU4
#cross-references GB:M22405
#experimental_source intestine
REFERENCE A61257
#authors Jany, B.H.; Gallup, M.W.; Yan, P.S.; Gum, J.R.; Kim, Y.S.;
Basbaum, C.B.
#journal J. Clin. Invest. (1991) 87:77-82
#title Human bronchus and intestine express the same mucin gene.
#accession A61257
#status not compared with conceptual translation
#molecule_type mRNA

```

```

##residues      'n',1925-1948,'TTS',1952-1954 ##label JAN
##experimental_source bronchus
REFERENCE      PQ0328
#authors      Xu, G.; Huan, L.; Khatri, I.; Sajjan, U.S.; McCool, D.; Wang,
              D.; Jones, C.; Forstner, G.; Forstner, J.
#journal      Biochem. Biophys. Res. Commun. (1992) 183:821-828
#title        Human intestinal mucin-like protein (MLP) is homologous with
              rat MLP in the C-terminal region, and is encoded by a gene
              on chromosome 11 p 15.5.
#cross-references MUID:92198477
#accession     PQ0328
##molecule_type mRNA
##residues      2328-2468 ##label XUG
##cross-references GB:M8523
##experimental_source small intestine
#accession     PQ0329
##molecule_type protein
##residues      2328-2342,'K',2344-2354 ##label XUG1
GENETICS
#gene          GDB:MUC2
#map_position  llp15.5-llp15.5
CLASSIFICATION #superfamily von Willebrand factor type C repeat homology
KEYWORDS       glycoprotein; tandem repeat
FEATURE        2766-2834      #domain von Willebrand factor type C repeat homology
              #label VWC
SUMMARY        #length 3020 #checksum 1599
Query Match    67.6%; Score 50; DB 13; Length 3020;
Best Local Similarity 75.0%; Pred. No. 6.56e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 2530 pspicql 2537
||| |||:
Qy 3 PSPACQLV 10

RESULT 10
ENTRY      SYPJCB      #type complete
TITLE      naringenin-chalcone synthase (EC 2.3.1.74) B - garden petunia
ALTERNATE_NAMES
ORGANISM    chalcone synthase
#formal_name Petunia x hybrida #common_name garden petunia
DATE        30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
              30-Jun-1993
ACCESSIONS  JS0311
REFERENCE    JS0308
#authors     Koes, R.E.; Spelt, C.E.; van den Elzen, P.J.M.; Mol, J.N.M.
#journal      Gene (1989) 81:245-257
#title        Cloning and molecular characterization of the chalcone
              synthase multigene family of Petunia hybrida.
#cross-references MUID:90034197
#accession    JS0311
##molecule_type DNA
##residues    1-392 ##label KOE
COMMENT       This enzyme plays a central role in the biosynthesis of all classes
              of flavonoids in plants.
CLASSIFICATION #superfamily chalcone synthase
KEYWORDS       acyltransferase; flavonoid biosynthesis
SUMMARY        #length 392 #molecular-weight 42956 #checksum 8913
Query Match    66.2%; Score 49; DB 1; Length 392;
Best Local Similarity 55.6%; Pred. No. 1.00e+01;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 140 mpgadcgly 148
|||: |||
Qy 2 LPSPACQLV 10

RESULT 11
ENTRY      A28755      #type fragment
TITLE      ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b

```

```

- Neurospora crassa mitochondrion (SGC3) (fragment)
#formal_name mitochondrion Neurospora crassa
15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change
03-May-1995
ACCESSIONS    A28755
REFERENCE      Burke, J.M.; Breitenberger, C.; Heckman, J.E.; Dujon, B.;
              RajBhandary, U.L.
#journal      J. Biol. Chem. (1984) 259:504-511
#title        Cytochrome b gene of Neurospora crassa mitochondria.
#cross-references MUID:84161957
#accession     A28755
##molecule_type DNA
##residues     1-568 ##label BUR
GENETICS
#genome        mitochondrion
#genetic_code  SGC3
#introns       347/1
CLASSIFICATION #superfamily plastoquinol--plastocyanin reductase 17K protein
              homology
KEYWORDS       electron transfer; heme; mitochondrion; oxidative
              phosphorylation; oxidoreductase; respiratory chain;
              transmembrane protein
FEATURE        406-524      #domain plastoquinol--plastocyanin reductase 17K protein
              homology #label 17K
SUMMARY        #length 568 #checksum 5132
Query Match    66.2%; Score 49; DB 12; Length 568;
Best Local Similarity 66.7%; Pred. No. 1.00e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 158 lrspctkly 166
|||: |||
Qy 2 LPSPACQLV 10

RESULT 12
ENTRY      S55474      #type complete
TITLE      Human giant larvae homolog - human
ALTERNATE_NAMES
ORGANISM      01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change
              01-Sep-1995
ACCESSIONS    S55474
REFERENCE      S55474
#authors      Wiemann, S.; Tommerup, N.; Celis, J.E.; Ansoorge, W.; Leffers,
              H.
#submission    submitted to the EMBL Data Library, May 1995
#description    A human homolog of the Drosophila 1(2) giant larvae tumor
              suppressor maps to 17q24-25.
#accession     S55474
##status       preliminary
##molecule_type mRNA
##residues     1-1015 ##label WIE
##cross-references EMBL:X87342
SUMMARY        #length 1015 #molecular-weight 112907 #checksum 9425
Query Match    66.2%; Score 49; DB 13; Length 1015;
Best Local Similarity 50.0%; Pred. No. 1.00e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 135 vlphscell 144
|||: |||:
Qy 1 VLPSPACQLV 10

RESULT 13
ENTRY      S25256      #type complete
TITLE      fur protein - Versinia pestis
ALTERNATE_NAMES
ORGANISM      12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
              12-Feb-1993
ACCESSIONS    S25256; S21004

```

```
REFERENCE S25256
#authors Stagg, T.M.; Perry, R.D.
#journal Mol. Microbiol. (1992) 6:2507-2516
#title Fur regulation in Yersinia species.
#accession S25256
##molecule_type DNA
##residues 1-148 ##label STA
##cross-references EMBL:212101
GENETICS
#gene fur
SUMMARY
#length 148 #molecular-weight 16721 #checksum 3240
Query Match 64.9%; Score 48; DB 9; Length 148;
Best Local Similarity 60.0%; Pred. No. 1.52e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 25 vlpnpachhv 34
|||:::|
QY 1 VLPSPACQLV 10

RESULT 14
ENTRY H64502 #type complete
TITLE hypothetical protein YPL009c homolog - Methanococcus
jannaschii
ORGANISM #formal_name Methanococcus jannaschii
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
13-Sep-1996
ACCESSIONS H64502
REFERENCE A64300
#authors Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
R.D.; Sutton, G.G.; Blake, J.A.; Fitzgerald, L.M.; Clayton,
R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann,
J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,
J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
#journal Science (1996) 273:1058-1073
#title Complete genome sequence of the methanogenic archaeon,
Methanococcus jannaschii.
#accession H64502
#status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-671 ##label BUL
##cross-references GB:L77117; TIGR:MJ1625; CDS_PID:g1511590
GENETICS
#map_position FOR1602213-1604228
SUMMARY #length 671 #molecular-weight 78615 #checksum 145
Query Match 64.9%; Score 48; DB 10; Length 671;
Best Local Similarity 50.0%; Pred. No. 1.52e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 660 vlpnpqgciv 669
|||:::|
QY 1 VLPSPACQLV 10

RESULT 15
ENTRY I51669 #type complete
TITLE tumor suppressor - African clawed frog
ORGANISM #formal_name Xenopus laevis #common_name African clawed frog
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
13-Sep-1996
ACCESSIONS I51669
REFERENCE I51668
#authors Pierceall, W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.;
Cho, K.R.; Fearon, E.R.
#journal Dev. Biol. (1994) 166:654-665
```

```
#title Expression of a homologue of the deleted in colorectal cancer
(DCC) gene in the nervous system of developing xenopus
embryos.
#cross-references MUID:95113183
#accession I51669
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues 1-1427 ##label PIE
##cross-references EMBL:U10986; MID:g606873; CDS_PID:g606874
GENETICS
#note gene name XDCCA
SUMMARY #length 1427 #molecular-weight 156533 #checksum 8558
Query Match 64.9%; Score 48; DB 13; Length 1427;
Best Local Similarity 71.4%; Pred. No. 1.52e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Db 1247 ilpsptc 1253
|||:::|
QY 1 VLPSPAC 7

Search completed: Tue Jun 10 11:27:17 1997
Job time : 11 secs.
```

\*\*\*\*\*

W A S E R E H (TM)

\*\*\*\*\*

Release 2.1d John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:26:41 1997; MasPar time 2.01 seconds  
Tabular output not generated. 105.754 Million cell updates/sec

Title: >US-08-231-565A-35  
Description: (1-10) from US08231565A.pap  
Perfect Score: 74  
Sequence: 1 VLSPSPACQLV 10

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 22.936; Variance 26.585; scale 0.863

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	74	100.0	668	7	PM17_HUMAN	1.97e-06
2	57	77.0	1056	6	MUC5_HUMAN	3.87e-02
3	50	67.6	353	6	MRP_SYN3	1.50e+00
4	50	67.6	2274	6	MUC2_HUMAN	1.50e+00
5	50	67.6	2704	4	G16E_PARP	1.50e+00
6	50	67.6	2715	4	G15G_PARP	1.50e+00
7	49	66.2	392	2	CHS2_PETH	2.47e+00
8	49	66.2	486	11	YAM8_SCHPO	2.47e+00
9	48	64.9	148	4	FUR_VERPE	4.03e+00
10	48	64.9	229	10	VP25_NPVOP	4.03e+00
11	48	64.9	1447	3	DCC_HUMAN	4.03e+00
12	47	63.5	329	2	CDK7_RAT	6.53e+00
13	47	63.5	346	2	CDK7_MOUSE	6.53e+00
14	47	63.5	346	2	CDK7_HUMAN	6.53e+00
15	47	63.5	549	5	LEM2_RAT	6.53e+00
16	47	63.5	612	5	LEM2_MOUSE	6.53e+00
17	47	63.5	1745	11	YL51_CABEL	6.53e+00
18	46	62.2	216	10	Y43_SCHPO	1.05e+01
19	46	62.2	223	10	UBL1_RAT	1.05e+01
20	46	62.2	291	4	GDBE_WHEAT	1.05e+01
21	46	62.2	308	6	MENA_ECOLI	1.05e+01
22	46	62.2	360	4	G3P_ATRNU	1.05e+01

23	46	62.2	1134	3	CYAL_BOVIN	ADENYLATE CYCLASE, TY	1.05e+01
24	45	60.8	72	9	SP21_HUMAN	SMALL PROLINE-RICH PR	1.68e+01
25	45	60.8	72	9	SP22_HUMAN	SMALL PROLINE-RICH PR	1.68e+01
26	45	60.8	72	9	SP2B_HUMAN	SMALL PROLINE-RICH PR	1.68e+01
27	45	60.8	72	9	SP23_HUMAN	SMALL PROLINE-RICH PR	1.68e+01
28	45	60.8	92	10	VNST_BUNSH	NONSTRUCTURAL PROTEIN	1.68e+01
29	45	60.8	298	2	CGIC_RAT	GL/S-SPECIFIC CYCLIN	1.68e+01
30	45	60.8	611	2	CALG_MOUSE	CALMEGIN PRECURSOR (M	1.68e+01
31	45	60.8	837	6	MUCL_RAT	INTESTINAL MUCIN-LIKE	1.68e+01
32	45	60.8	978	7	PAF2_RAT	PEROXISOME ASSEMBLY F	1.68e+01
33	45	60.8	1266	9	SYL_HUMAN	ISOLEUCYL-TRNA SYNTH	1.68e+01
34	45	60.8	1290	7	PIP4_HUMAN	1-PHOSPHATIDYLINOSITO	1.68e+01
35	45	60.8	1290	7	PIP4_RAT	1-PHOSPHATIDYLINOSITO	1.68e+01
36	45	60.8	1291	7	PIP4_BOVIN	1-PHOSPHATIDYLINOSITO	1.68e+01
37	45	60.8	3343	11	YOG7_CABEL	HYPOTHETICAL 375.7 KD	1.68e+01
38	44	59.5	160	9	THY1_CHICK	THY-1 MEMBRANE GLYCOP	2.65e+01
39	44	59.5	348	5	HM14_CABEL	HOMEOBOX PROTEIN CEH-	2.65e+01
40	44	59.5	514	11	YHK8_YEAST	HYPOTHETICAL 57.8 KD	2.65e+01
41	44	59.5	641	4	HEMA_INCKY	HEMAGGLUTININ.	2.65e+01
42	44	59.5	642	4	HEMA_INCP1	HEMAGGLUTININ PRECURS	2.65e+01
43	44	59.5	704	5	HS85_TRYCR	HEAT SHOCK LIKE 85 KD	2.65e+01
44	44	59.5	1018	11	YKC2_CABEL	HYPOTHETICAL 115.2 KD	2.65e+01
45	44	59.5	1080	11	YHR9_YEAST	HYPOTHETICAL 123.9 KD	2.65e+01

ALIGNMENTS

RESULT	1	PM17_HUMAN	STANDARD;	PRT;	668 AA.
AC	P40967;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	KWON B.S., CHINTAMANENI C., KOZAK C.A., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,				
RA	KIM K.-K.,				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).				
CC	-!- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	-!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; G190106; -				
DR	MIM; 155550; -				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23		POTENTIAL.
FT	CHAIN	24	668		PMEL 17 PROTEIN.
FT	TRANSMEM	575	595		POTENTIAL.
FT	TRANSMEM	603	623		POTENTIAL.
FT	DOMAIN	217	307		PKD.
FT	DOMAIN	315	444		10 X 13 AA TANDEM REPEATS.
FT	REPEAT	315	327	1.	
FT	REPEAT	328	340	2.	
FT	REPEAT	341	353	3.	
FT	REPEAT	354	366	4.	
FT	REPEAT	367	379	5.	
FT	REPEAT	380	392	6.	
FT	REPEAT	393	405	7.	
FT	REPEAT	406	418	8.	
FT	REPEAT	419	431	9.	
FT	REPEAT	432	444	10.	
FT	CARBOHYD	81	81		POTENTIAL.
FT	CARBOHYD	106	106		POTENTIAL.
FT	CARBOHYD	111	111		POTENTIAL.
FT	CARBOHYD	321	321		POTENTIAL.
FT	CARBOHYD	568	568		POTENTIAL.
FT	SEQUENCE	668 AA;	70992 MW;		6E8E1AFO CRC32;

Query Match 100.0%; Score 74; DB 7; Length 668;  
Best Local Similarity 100.0%; Pred. No. 1.97e-06;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 544 vlspacqlv 553  
QY 1 VLSPACQLV 10

RESULT 2  
ID MUC5\_HUMAN STANDARD; PRT: 1056 AA.  
AC P98088;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE TRACHEOBRONCHIAL MUCIN (TBM) (MAJOR AIRWAY GLYCOPROTEIN) (FRAGMENT).  
GN MUC5AC OR MUC5.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-44; 65-76 AND 493-496.  
RC TISSUE=LUNG;  
RX MEDLINE; 94230376.  
RA MEERZAMAN D., CHARLES P., DASKAL E., POLYMERPOULOS M.H.,  
RA MARTIN B.M., ROSE M.C.;  
RA J. BIOL. CHEM. 269:12932-12939(1994).  
RL  
CC -!- FUNCTION: MAJOR AIRWAY GLYCOPROTEIN.  
CC -!- TISSUE SPECIFICITY: TRACHEOBRONCHIAL MUCOSAE (NOT SOLELY).  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SUBUNIT: MULTIMERIC.  
CC -!- SIMILARITY: TO VARIOUS MUCINS.  
CC -!- SIMILARITY: CONTAINS A VWFC DOMAIN.  
CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).  
DR EMBL; U06711; G488073; -.  
DR MIM; 158373; -.  
KW REPEAT: GLYCOPROTEIN.  
FT NON\_TER 1  
FT DOMAIN 496 520 5 X 5 AA TANDEM REPEATS OF T-T-V-G-P/S.  
FT REPEAT 496 500 1.  
FT REPEAT 501 505 2.  
FT REPEAT 506 510 3.  
FT REPEAT 511 515 4.  
FT REPEAT 516 520 5.  
FT DOMAIN 661 730 VWFC.  
FT DOMAIN 917 1005 CTCK.  
FT DISULFID 917 967 BY SIMILARITY.  
FT DISULFID 942 981 BY SIMILARITY.  
FT DISULFID 946 997 BY SIMILARITY.  
FT DISULFID 964 999 BY SIMILARITY.  
FT DISULFID 966 1004 BY SIMILARITY.  
FT CARBOHYD 256 256 POTENTIAL.  
FT CARBOHYD 443 443 POTENTIAL.  
FT CARBOHYD 479 479 POTENTIAL.  
FT CARBOHYD 621 621 POTENTIAL.  
FT CARBOHYD 732 732 POTENTIAL.  
FT CARBOHYD 762 762 POTENTIAL.  
FT CARBOHYD 771 771 POTENTIAL.  
FT CARBOHYD 913 913 POTENTIAL.  
FT CARBOHYD 976 976 POTENTIAL.  
FT SIMILAR 1 121 TO N-TERMINAL CYSTEINE-RICH DOMAIN OF  
FT SIMILAR 113 1024 HUMAN INTESTINAL MUCIN (MUC2).  
FT SIMILAR 215 729 TO C-TERMINAL CYSTEINE-RICH DOMAIN OF  
FT SIMILAR 723 795 HUMAN INTESTINAL MUCIN (MUC2).  
FT SIMILAR 777 1009 TO RAT MUCIN-LIKE PROTEIN.  
FT SEQUENCE 1056 AA; 113043 MW; C7D1BFA CRC32; TO BOVINE SUBMAXILLARY MUCIN.  
TO PORCINE SUBMAXILLARY MUCIN.

Query Match 77.0%; Score 57; DB 6; Length 1056;  
Best Local Similarity 66.7%; Pred. No. 3.87e-02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 528 lpsplchli 536  
QY 2 LPSPACQLV 10

RESULT 3  
ID MRP\_SYNY3 STANDARD; PRT: 353 AA.  
AC P5383;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE MRP PROTEIN HOMOLOG.  
GN MRP OR SLR0067.  
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).  
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;  
OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96127529.  
RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,  
RA SUGIURA M., TABATA S.;  
RA DNA RES. 2:153-166(1995).  
RL  
CC -!- FUNCTION: NOT KNOWN.  
CC -!- SIMILARITY: BELONGS TO THE MRP/NBP35 FAMILY OF ATP-BINDING  
CC PROTEINS.  
CC EMBL; D64001; G1001155; -.  
DR ATP-BINDING.  
KW NP\_BIND 105 112 ATP (POTENTIAL).  
FT NP\_BIND 105 112 ATP (POTENTIAL).  
SQ SEQUENCE 353 AA; 37131 MW; 25125F58 CRC32;

Query Match 67.6%; Score 50; DB 6; Length 353;  
Best Local Similarity 66.7%; Pred. No. 1.50e+00;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 44 vltpacpl 52  
QY 1 VLSPACOL 9

RESULT 4  
ID MUC2\_HUMAN STANDARD; PRT: 2274 AA.  
AC Q02817;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE MUCIN 2 (INTESTINAL MUCIN 2) (FRAGMENTS).  
GN MUC2 OR SMUC.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE OF 1-1270 AND 1291-2274 FROM N.A.  
RC TISSUE=COLON;  
RX MEDLINE; 93016075.  
RA GUM J.R. JR., HICKS J.W., TORIBARA N.W., ROTHE E.-M., LAGACE R.E.,  
RA KIM Y.S.;  
RA J. BIOL. CHEM. 267:21375-21383(1992).  
RN [2]  
RP SEQUENCE OF 718-1290 FROM N.A.  
RX MEDLINE; 91358717.  
RA TORIBARA N.W., GUM J.R. JR., CULHANE P.J., LAGACE R.E., HICKS J.W.,  
RA PETERSEN G.M., KIM Y.S.;  
RA J. CLIN. INVEST. 88:1005-1013(1991).  
CC -!- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND  
CC -!- SUBCELLULAR LOCATION: SECRETED. OTHER MUCUS MEMBRANE-CONTAINING ORGANS.  
CC -!- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR  
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).  
CC -!- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,  
CC BRONCHUS, CERVIX AND GALL BLADDER.  
CC -!- SUBUNIT: MULTIMERIC.  
CC -!- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND

CC VARIES AMONG DIFFERENT ALLELES.  
CC -!- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT  
CC OF SILKWORM HEMOCYTIN.  
CC -!- SIMILARITY: CONTAINS 2 VWFC DOMAINS.  
CC -!- SIMILARITY: CONTAINS A C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).  
DR EMBL; M74027; G188864; -.  
DR EMBL; M94131; G186396; -.  
DR EMBL; M94132; G186398; -.  
DR HSSP; P15445; LP5H.  
DR MIN; 158370; -.  
DR PROSITE; PS00022; EGF.  
KW GLYCOPROTEIN; REPEAT; POLYMORPHISM.  
FT NON\_TER 1 1  
FT DOMAIN 776 1122 APPROXIMATE REPEATS.  
FT REPEAT 776 791 1.  
FT REPEAT 792 807 2.  
FT REPEAT 808 823 3.  
FT REPEAT 824 839 4.  
FT REPEAT 840 846 5.  
FT REPEAT 847 853 6.  
FT REPEAT 854 869 7A.  
FT REPEAT 870 892 7B.  
FT REPEAT 893 908 8A.  
FT REPEAT 909 931 8B.  
FT REPEAT 932 947 9A.  
FT REPEAT 948 971 9B.  
FT REPEAT 972 987 10A.  
FT REPEAT 988 1010 10B.  
FT REPEAT 1011 1026 11A.  
FT REPEAT 1027 1050 11B.  
FT REPEAT 1051 1058 12.  
FT REPEAT 1059 1074 13.  
FT REPEAT 1075 1090 14.  
FT REPEAT 1091 1106 15.  
FT REPEAT 1107 1122 16.  
FT DOMAIN 1310 1981 VWFC 1.  
FT DOMAIN 2019 2086 VWFC 2.  
FT DOMAIN 2170 2255 CTCK.  
FT DISULFID 2170 2217 BY SIMILARITY.  
FT DISULFID 2184 2231 BY SIMILARITY.  
FT DISULFID 2193 2247 BY SIMILARITY.  
FT DISULFID 2197 2249 BY SIMILARITY.  
FT DISULFID ? 2254 BY SIMILARITY.  
FT CARBOHYD 1434 1434 POTENTIAL.  
FT CARBOHYD 1446 1446 POTENTIAL.  
FT CARBOHYD 1457 1457 POTENTIAL.  
FT CARBOHYD 1468 1468 POTENTIAL.  
FT CARBOHYD 1517 1517 POTENTIAL.  
FT CARBOHYD 1533 1533 POTENTIAL.  
FT CARBOHYD 1597 1597 POTENTIAL.  
FT CARBOHYD 1711 1711 POTENTIAL.  
FT CARBOHYD 1722 1722 POTENTIAL.  
FT CARBOHYD 1847 1847 POTENTIAL.  
FT CARBOHYD 1882 1882 POTENTIAL.  
FT CARBOHYD 1976 1976 POTENTIAL.  
FT CARBOHYD 1983 1983 POTENTIAL.  
FT CARBOHYD 2050 2050 POTENTIAL.  
FT CARBOHYD 2065 2065 POTENTIAL.  
FT CARBOHYD 2113 2113 POTENTIAL.  
FT CARBOHYD 2133 2133 POTENTIAL.  
FT CARBOHYD 2164 2164 POTENTIAL.  
FT CONFLICT 726 726 H -> L (IN REF. 2).  
FT CONFLICT 787 787 T -> S (IN REF. 2).  
FT CONFLICT 824 824 L -> P (IN REF. 2).  
FT CONFLICT 879 879 M -> T (IN REF. 2).  
FT NON\_CONS 1290 1291  
SQ SEQUENCE 2274 AA; 244013 MW; A05018ED CRC32;  
Query Match 67.6%; Score 50; DB 6; Length 2274;  
Best Local Similarity 75.0%; Pred. No. 1.50e+00;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 1784 pspicqli 1791

QY 3 PSPACOLV 10  
RESULT 5  
ID G168.PARPR STANDARD; PRT; 2704 AA.  
AC P17053;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-AUG-1990 (REL. 15, LAST ANNOTATION UPDATE)  
DE 168G SURFACE PROTEIN PRECURSOR.  
GN 168G.  
OS PARAMECIUM PRIMAURELIA.  
OC EUKARYOTA; PROTOZOA; CILIOPHORA; CILIATA; HOLOTRICHA; HYMENOSTOMATIDA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE; 90172419.  
RA PRAT A.;  
RL J. MOL. BIOL. 211:521-535(1990).  
CC -!- FUNCTION: THIS PROTEIN IS THE SURFACE ANTIGEN OR IMMOBILIZATION  
CC ANTIGEN OF PARAMECIUM PRIMAURELIA.  
CC -!- IT HAS INTERNAL HOMOLOGIES AND A HIGHLY PERIODIC STRUCTURE WITH  
CC 37 PERIODS OF ABOUT 75 RESIDUES, EACH PERIOD CONTAINING 8  
CC CYSTEINES, EXCEPT FOR FOUR HALF PERIODS. A VARIABLE PART OF 475  
CC RESIDUES COMPRISES 4 ALMOST IDENTICAL PERIODS IN THE MIDDLE OF THE  
CC PROTEIN.  
CC -!- EXPRESSION OF G PROTEIN OCCURS AT LOW TEMPERATURES (14-32  
CC DEGREES CELSIUS).  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
CC -!- SIMILARITY: 98% TO THE ALLELIC FORM 156G PROTEIN (P13837) IN  
CC PERIODIC STRUCTURE AND 80% IN THE VARIABLE DOMAIN IN THE MIDDLE  
CC OF THE PROTEIN.  
DR EMBL; X52133; G578473; -.  
DR PIR; S09118; S09118.  
DR HSSP; P00766; LCHG.  
KW SIGNAL; REPEAT; ANTIGEN; MEMBRANE; GPI-ANCHOR.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 2704 168G SURFACE PROTEIN.  
FT DOMAIN 106 2560 37 X 75 AA APPROXIMATE REPEATS.  
FT DOMAIN 1060 1424 VARIABLE DOMAIN, COMPRISES 4 ALMOST  
FT IDENTICAL REPEATS.  
SQ SEQUENCE 2704 AA; 278775 MW; 5B49B158 CRC32;  
Query Match 67.6%; Score 50; DB 4; Length 2704;  
Best Local Similarity 85.7%; Pred. No. 1.50e+00;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Db 1876 lpapacq 1882  
QY 2 LFSPACQ 8  
RESULT 6  
ID G156.PARPR STANDARD; PRT; 2715 AA.  
AC P13837;  
DT 01-JAN-1990 (REL. 13, CREATED)  
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
DT 01-AUG-1990 (REL. 15, LAST ANNOTATION UPDATE)  
DE 156G SURFACE PROTEIN PRECURSOR.  
GN 156G.  
OS PARAMECIUM PRIMAURELIA.  
OC EUKARYOTA; PROTOZOA; CILIOPHORA; CILIATA; HOLOTRICHA; HYMENOSTOMATIDA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=156;  
RX MEDLINE; 87060934.  
RA PRAT A.; RATINKA M.; CARON F.; MEYER E.;  
RL J. MOL. BIOL. 189:47-60(1986).  
CC -!- FUNCTION: THIS PROTEIN IS THE SURFACE ANTIGEN OR IMMOBILIZATION  
CC ANTIGEN OF PARAMECIUM PRIMAURELIA.  
CC -!- IT HAS INTERNAL HOMOLOGIES AND A HIGHLY PERIODIC STRUCTURE WITH  
CC 37 PERIODS OF ABOUT 75 RESIDUES, EACH PERIOD CONTAINING 8



CC CYSTEINES, EXCEPT FOR FOUR HALF PERIODS. A VARIABLE PART OF 475  
 CC RESIDUES COMPRISES 4 ALMOST IDENTICAL PERIODS IN THE MIDDLE OF THE  
 CC PROTEIN.  
 CC -1- EXPRESSION OF G PROTEIN OCCURS AT LOW TEMPERATURES (14-32  
 CC DEGREES CELSIUS)  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
 CC -1- SIMILARITY: 98% TO THE ALLELIC FORM 168G PROTEIN (P17053) IN  
 CC PERIODIC STRUCTURE AND 80% IN VARIABLE DOMAIN IN THE MIDDLE OF  
 CC THE PROTEIN.  
 CC EMBL: X03882; G578475; -.  
 CC PIR: A23475; A23475.  
 CC HSP: P06620; IINA.  
 CC SIGNAL; REPEAT; ANTIGEN; MEMBRANE; GPI-ANCHOR.  
 CC FT SIGNAL 1 20 POTENTIAL.  
 CC FT CHAIN 21 2715 156G SURFACE PROTEIN.  
 CC FT DOMAIN 106 2560 37 X 75 AA APPROXIMATE REPEATS.  
 CC FT SIMILAR 1 222 88% TO PARAMECIUM TETRAURELIA A  
 CC PROTEIN.  
 CC SQ SEQUENCE 2715 AA; 279551 MW; 64FF65AA CRC32;  
 Query Match 67.6%; Score 50; DB 4; Length 2715;  
 Best Local Similarity 85.7%; Pred. No. 1.50e+00;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 1887 lpapacq 1893  
 QY 2 LPSPACQ 8  
 RESULT 7  
 ID CHSB\_PETHY STANDARD; PRT; 392 AA.  
 AC P22924;  
 DT 01-AUG-1991 (REL. 19, CREATED)  
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
 DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
 DE CHALCONE SYNTHASE B (EC 2.3.1.74) (NARINGENIN-CHALCONE SYNTHASE B).  
 GN CHSB.  
 OS PETUNIA HYBRIDA (PETUNIA).  
 OS EUKARYOTA; PLANTA; EMERYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;  
 OS SOLANALES; SOLANACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-CV, VIOLET 30; TISSUE-LEAF;  
 RX MEDLINE: 90034197.  
 RA KOES R.E., SPELT C.E., VAN DEN ELZEN P.J.M., MOL J.N.M.;  
 RL GENE 81:245-257(1989).  
 CC -1- CATALYTIC ACTIVITY: 3 MALONYL-COA + 4-COUMAROYL-COA = 4 COA +  
 CC NARINGENIN-CHALCONE + 3 CO(2).  
 CC -1- PATHWAY: THIS PROTEIN IS AN ENZYME OF CENTRAL IMPORTANCE IN THE  
 CC BIOSYNTHESIS PATHWAY OF ALL CLASSES OF FLAVONOIDS IN PLANTS.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVEL IN SEEDLINGS AFTER  
 CC ILLUMINATION WITH UV LIGHT. NO EXPRESSION IN FLOWERS OR TISSUE  
 CC CULTURE.  
 CC -1- SIMILARITY: TO OTHER PLANTS CHALCONE AND STILBENE SYNTHASES.  
 CC EMBL: X14592; G20527; -.  
 CC PIR: JS0311; SVPJCB.  
 CC PROSITE: PS00441; CHALCONE\_SYNTH.  
 CC FLAVONOID BIOSYNTHESIS; TRANSFERASE; ACYLTRANSFERASE;  
 CC MULTIGENE FAMILY.  
 CC ACT\_SITE 167 167 BY SIMILARITY.  
 CC SQ SEQUENCE 392 AA; 42956 MW; 4F982155 CRC32;  
 Query Match 66.2%; Score 49; DB 2; Length 392;  
 Best Local Similarity 55.6%; Pred. No. 2.47e+00;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 140 mpegdcqlv 148  
 QY 2 LPSPACQLV 10  
 RESULT 8  
 ID YAM8\_SCHPO STANDARD; PRT; 486 AA.

AC Q10063;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 53.9 KD PROTEIN C1F5.08C IN CHROMOSOME I.  
 GN SPAC1F5.08C.  
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
 OS EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA GENTLES S.; ODELL C.; CHURCHER C.M.; BARRELL B.G.; RAJANDREAM M.A.;  
 RA WALSH S.V.;  
 RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- SIMILARITY: SOME, TO YEAST MID1.  
 DR EMBL; Z68136; E212001; -.  
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE; GLYCOPROTEIN.  
 FT TRANSMEM 12 32 POTENTIAL.  
 FT TRANSMEM 465 485 POTENTIAL.  
 FT CARBOHYD 33 33 POTENTIAL.  
 FT CARBOHYD 49 49 POTENTIAL.  
 FT CARBOHYD 59 59 POTENTIAL.  
 FT CARBOHYD 82 82 POTENTIAL.  
 FT CARBOHYD 93 93 POTENTIAL.  
 FT CARBOHYD 160 160 POTENTIAL.  
 FT CARBOHYD 181 181 POTENTIAL.  
 FT CARBOHYD 185 185 POTENTIAL.  
 FT CARBOHYD 201 201 POTENTIAL.  
 FT CARBOHYD 255 255 POTENTIAL.  
 FT CARBOHYD 307 307 POTENTIAL.  
 FT CARBOHYD 330 330 POTENTIAL.  
 FT CARBOHYD 341 341 POTENTIAL.  
 FT CARBOHYD 353 353 POTENTIAL.  
 FT CARBOHYD 379 379 POTENTIAL.  
 FT CARBOHYD 445 445 POTENTIAL.  
 FT CARBOHYD 468 468 POTENTIAL.  
 SQ SEQUENCE 486 AA; 53888 MW; 098328CD CRC32;  
 Query Match 66.2%; Score 49; DB 11; Length 486;  
 Best Local Similarity 85.7%; Pred. No. 2.47e+00;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 285 spacqli 291  
 QY 4 SPACQLV 10  
 RESULT 9  
 ID FUR\_YERPE STANDARD; PRT; 148 AA.  
 AC P33086;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE FERRIC UPTAKE REGULATION PROTEIN.  
 GN FUR.  
 OS YERSINIA PESTIS.  
 OS PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC ENTEROBACTERIACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-KIM 6;  
 RX MEDLINE: 93023880.  
 RA STAGGS T.M., PERRY R.D.;  
 RL MOL. MICROBIOL. 6:2507-2516(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-KIM 6;  
 RA BEARDEN S.W.;  
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -1- FUNCTION: FUR ACTS AS A REPRESSOR, EMPLOYING FE(2+) AS A COFACTOR  
 CC TO BIND THE OPERATOR OF THE IRON TRANSPORT OPERON.  
 CC -1- SIMILARITY: TO OTHER BACTERIA FUR AND TO E.COLI YJBK.  
 DR EMBL; Z12101; G871776; -.

DR PIR: S25256; S25256.  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; IRON.  
 FT DOMAIN 86 90 HIS-RICH.  
 FT CONFLICT 58 60 VLN -> CSE (IN REF. 1).  
 SQ SEQUENCE 148 AA; 16728 MW; 90724A73 CRC32;  
 Query Match 64.9%; Score 48; DB 4; Length 148;  
 Best Local Similarity 60.0%; Pred. No. 4.03e+00;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Db 25 vlpnpachv 34  
 ||:||:|  
 QY 1 VLSPACQLV 10

RESULT 10  
 ID VP25\_NPVOP STANDARD; PRT; 229 AA.  
 AC Q06906;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE P25 PROTEIN.  
 GN P25.  
 OS ORGYIA PSEUDOTSUGATA MULTICAPSID POLYHEDROSIS VIRUS (OPMNPV).  
 CC VIRIDAE; DS-DNA ENVELOPED VIRUSES; BACULOVIRIDAE; EUBACULOVIRINAE.  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE: 93331711.  
 RX RUSSELL R.L.O., ROHRMANN G.F.;  
 RA VIROLOGY 195:532-540(1993).  
 RL  
 CC -1- FUNCTION: COMPONENT OF THE POLYHEDRA ENVELOPE.  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE ENVELOPES OF VIRIONS  
 CC PRESENT IN THE NUCLEUS OF INFECTED CELLS, WITH THE ENVELOPES OF  
 CC VIRIONS IN THE PROCESS OF BEING OCCLUDED AND FULLY OCCLUDED  
 CC VIRIONS.  
 CC  
 CC -1- DEVELOPMENTAL STAGE: FIRST APPEARS AFTER 18 HRS OF INFECTION.  
 CC -1- SIMILARITY: TO THE CORRESPONDING PROTEIN OF ACMNPV.  
 DR EMBL: D13768; G222192; -.  
 KW NUCLEAR PROTEIN.  
 SQ SEQUENCE 229 AA; 25325 MW; 3FBC4340 CRC32;

Query Match 64.9%; Score 48; DB 10; Length 229;  
 Best Local Similarity 66.7%; Pred. No. 4.03e+00;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 185 lpsactlv 193  
 ||:||||  
 QY 2 LPSPACQLV 10

RESULT 11  
 ID DCC\_HUMAN STANDARD; PRT; 1447 AA.  
 AC P43146;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE TUMOR SUPPRESSOR PROTEIN DCC PRECURSOR (COLORECTAL CANCER SUPPRESSOR).  
 GN DCC.  
 OS HOMO SAPIENS (HUMAN).  
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 CC EUTHERIA; PRIMATES.  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE: 95011532.  
 RX HEDRICK L., CHO K.R., FEARON E.R., WU T.-C., KINZLER K.W.,  
 RA VOGELSTEIN B.;  
 RA GENES DEV. 8:1174-1183(1994).  
 RL  
 RN SEQUENCE 1-750 FROM N.A.  
 RP MEDLINE: 90100559.  
 RX FEARON E.R., CHO K.R., NIGRO J.M., KERN S.E., SIMONS J.W.,  
 RA RUPPERT J.M., HAMILTON S.R., PREISINGER A.C., THOMAS G., KINZLER K.W.,  
 RA VOGELSTEIN B.;

RL SCIENCE 247:49-56(1990).  
 [3]  
 RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBELD EXONS).  
 RX MEDLINE: 91121517.  
 RA NIGRO J.M., CHO K.R., FEARON E.R., KERN S.E., RUPPERT J.M.,  
 RA OLINER J.D., KINZLER K.W., VOGELSTEIN B.;  
 RA CELL 64:607-613(1991).  
 [4]  
 RN GENE STRUCTURE, AND VARIANTS CARCINOMA HIS-1375.  
 RP MEDLINE: 94245241.  
 RX CHO K.R., OLINER J.D., SIMONS J.W., HEDRICK L., FEARON E.R.,  
 RA PREISINGER A.C., HEDGE P., SILVERMAN G.A., VOGELSTEIN B.;  
 RL GENOMICS 19:525-531(1994).  
 [5]  
 RN VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.  
 RP MEDLINE: 94243823.  
 RX MIYAKE S., NAGAI K., YOSHINO K., OTO M., ENDO M., YUASA Y.;  
 RL CANCER RES. 54:3007-3010(1994).  
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL  
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC FOUR C2-LIKE DOMAINS FOLLOWED BY SIX FIBRONECTIN TYPE III-LIKE  
 CC DOMAINS.  
 CC -1- DISEASE: COLORECTAL TUMORS THAT LOST THEIR CAPACITY TO  
 CC DIFFERENTIATE INTO MUCUS PRODUCING CELLS UNIFORMLY LACK DCC  
 CC EXPRESSION. INACTIVATION OF DCC DUE TO ALLELIC DELETION AND/OR  
 CC POINT MUTATIONS MAY CAUSE BOTH LYMPHATIC AND HEMATOGENOUS  
 CC METASTASIS OF OESOPHAGEAL SQUAMOUS CELL CARCINOMAS.  
 DR EMBL: X76132; G453210; -.  
 DR EMBL: M32292; G181493; -.  
 DR EMBL: M32286; G441146; -.  
 DR EMBL: M32288; G441147; ALT\_SEQ.  
 DR EMBL: M32290; G533248; -.  
 DR EMBL: M32290; G533248; -.  
 DR EMBL: M63696; G181498; -.  
 DR EMBL: M63700; G533249; -.  
 DR EMBL: M63702; G533250; -.  
 DR EMBL: M63718; G533251; -.  
 DR EMBL: M63698; G181506; -.  
 DR PIR: A54100; A54100.  
 DR PIR: A40098; A40098.  
 DR PIR: A38442; A38442.  
 DR MIM: 120470; -.  
 KW GLYCOPROTEIN; IMMUNOGLOBULIN FOLD; TRANSMEMBRANE; SIGNAL;  
 FT ANTI-ONCOGENE; DISEASE MUTATION; POLYMORPHISM.  
 FT SIGNAL 1 25  
 FT CHAIN 26 1447  
 FT DOMAIN 26 1097  
 FT TRANSMEM 1098 1122  
 FT DOMAIN 1123 1447  
 FT DOMAIN 54 124  
 FT DOMAIN 154 219  
 FT DOMAIN 254 317  
 FT DOMAIN 345 407  
 FT DOMAIN 426 522  
 FT DOMAIN 525 618  
 FT DOMAIN 619 716  
 FT DOMAIN 722 816  
 FT DOMAIN 840 940  
 FT DOMAIN 941 1042  
 FT DISULFID 61 117  
 FT DISULFID 161 212  
 FT DISULFID 261 310  
 FT DISULFID 352 400  
 FT CARBOHYD 94 94  
 FT CARBOHYD 299 299  
 FT CARBOHYD 318 318  
 FT CARBOHYD 478 478  
 FT CARBOHYD 628 628  
 FT CARBOHYD 702 702  
 FT VARIANT 168 168  
 FT VARIANT 201 201  
 M -> T (IN OESOPHAGEAL CARCINOMA).  
 R -> G.

```

FT VARIANT 1375 1375 P -> H (IN A COLORECTAL CARCINOMA).
FT CONFLICT 138 MISSING (IN REF. 3).
FT CONFLICT 233 329 MISSING (IN REF. 3).
FT CONFLICT 421 421 MISSING (IN REF. 3).
SQ SEQUENCE 1447 AA; 158456 MW; 0D66434A CRC32;

Query Match 64.9%; Score 48; DB 3; Length 1447;
Best Local Similarity 71.4%; Pred. No. 4.03e+00; Indels 0; Gaps 0;
Matches 5; Conservative 2; Mismatches 0;

Db 1264 ilpsptc 1270
QY 1 VLPSPAC 7

RESULT 12
ID CDK7_RAT STANDARD; PRT; 329 AA.
AC P51952;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CELL DIVISION PROTEIN KINASE 7 (EC 2.7.1.-) (CDK-ACTIVATING KINASE)
DE (CAK) (39 KD PROTEIN KINASE) (P39 MO15) (FRAGMENT).
GN CDK7 OR CAK1.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-TESTIS;
RA WU L., HALL F.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CATALYTIC PART OF CAK WHICH ACTIVATES CYCLIN-ASSOCIATED
CC CDK2/CDK4 BY THREONINE PHOSPHORYLATION
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT T-170 IS REQUIRED FOR
CC ENZYME ACTIVITY (BY SIMILARITY).
CC -!- SUBUNIT: PROBABLY ASSOCIATES WITH CYCLIN H AND MAT1 TO FORM A
CC MULTIMERIC ACTIVE ENZYME (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CDC2/CDC28 SUBFAMILY OF SER/THR
CC PROTEIN KINASES.
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION, A STOP CODON WAS READ
CC THROUGH IN POSITION 313 TO MAXIMIZE SIMILARITIES WITH OTHER
CC SPECIES CDK7.
DR EMBL: X819509; G619509; ALT SEQ.
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; MEIOSIS;
KW PHOSPHORYLATION; CELL CYCLE; CELL DIVISION; NUCLEAR PROTEIN.
FT NON_TER 1 1
FT DOMAIN 1 4 287 PROTEIN KINASE (BY SIMILARITY).
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 129 129 BY SIMILARITY.
FT MOD_RES 136 136 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 162 162 PHOSPHORYLATION (BY SIMILARITY).
FT NON_TER 329 329
SQ SEQUENCE 329 AA; 37164 MW; 6751E918 CRC32;

Query Match 63.5%; Score 47; DB 2; Length 329;
Best Local Similarity 71.4%; Pred. No. 6.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 293 ptpgcl 299
QY 3 PSPACOL 9

RESULT 13
ID CDK7_MOUSE STANDARD; PRT; 346 AA.
AC Q03147;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CELL DIVISION PROTEIN KINASE 7 (EC 2.7.1.-) (CDK-ACTIVATING KINASE)

```

```

DE (CAK) (39 KD PROTEIN KINASE) (P39 MO15) (PROTEIN-TYROSINE KINASE MPK-
DE 7) (CR4 PROTEIN KINASE).
GN CDK7 OR MO15 OR CRK4 OR MPK-7.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA; TISSUE-BONE MARROW;
RX MEDLINE; 95047496.
RA STEPANOVA L.Y., ERSHLER M., BELYAVSKY A.V.;
RL GENE 149:321-324(1994).
RN [3]
RP SEQUENCE OF 139-154 FROM N.A.
RC STRAIN-C57BL; TISSUE-EMBRYONIC BRAIN;
RX MEDLINE; 93096484.
RA GILARDI-HEBENSTREIT P., NIETO M.A., FRAIN M., MATTEI M.-G.,
RA CHESTIER A., WILKINSON D.G., CHARNAY P.;
RL ONCOGENE 7:2499-2506(1992).
RN [4]
RP SEQUENCE OF 141-176 FROM N.A.
RC STRAIN-CBA; TISSUE-BONE MARROW;
RX MEDLINE; 93185941.
RA ERSHLER M.A., NAGORSKAYA T.V., VISSER J.W.M., BELYAVSKY A.V.;
RL GENE 124:305-306(1993).
RN [5]
RP SEQUENCE OF 141-176 FROM N.A.
RC STRAIN-CBA; TISSUE-BONE MARROW;
RX MEDLINE; 93092802.
RA ERSHLER M.A., NAGORSKAYA T.V., VISSER J.W.M., BELYAVSKII A.V.;
RL DOKL. AKAD. NAUK SSSR 324:893-897(1992).
CC -!- FUNCTION: CATALYTIC PART OF CAK WHICH ACTIVATES CYCLIN-ASSOCIATED
CC CDK2/CDK4 BY THREONINE PHOSPHORYLATION. EXPRESSION IS
CC INVARIANT DURING CELL CYCLE.
CC -!- ENZYME REGULATION: PHOSPHORYLATION AT T-170 IS REQUIRED FOR
CC ENZYME ACTIVITY (BY SIMILARITY).
CC -!- SUBUNIT: PROBABLY ASSOCIATES WITH CYCLIN H AND MAT1 TO FORM A
CC MULTIMERIC ACTIVE ENZYME (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN
CC KINASES. BELONGS TO THE CDC2/CDC28 SUBFAMILY.
DR EMBL: U11822; G562024; -.
DR EMBL: X74145; G395048; -.
DR EMBL: X57239; G53200; -.
DR EMBL: X65070; G296751; -.
DR EMBL: S55243; G261417; -.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM.
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW PHOSPHORYLATION; CELL CYCLE; CELL DIVISION; NUCLEAR PROTEIN.
FT DOMAIN 12 295
FT NP_BIND 18 26 ATP (BY SIMILARITY).
FT BINDING 41 41 ATP (BY SIMILARITY).
FT ACT_SITE 137 137 BY SIMILARITY.
FT MOD_RES 164 164 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 170 170 PHOSPHORYLATION (BY SIMILARITY).
FT CONFLICT 11 12 RY -> HN (IN REF. 2).
FT CONFLICT 20 20 E -> R (IN REF. 2).
FT CONFLICT 100 100 V -> L (IN REF. 2).
FT CONFLICT 104 104 D -> H (IN REF. 2).
FT CONFLICT 130 131 QH -> HN (IN REF. 2).
SQ SEQUENCE 346 AA; 38968 MW; 75A1FCH1 CRC32;

Query Match 63.5%; Score 47; DB 2; Length 346;
Best Local Similarity 71.4%; Pred. No. 6.53e+00;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

Db 301 ptpgqcl 307  
1:1:111  
QY 3 PSPACQL 9

```
RESULT 14
ID CDK7_HUMAN STANDARD; PRT; 346 AA.
AC P50613;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CELL DIVISION PROTEIN KINASE 7 (EC 2.7.1.-) (CDK-ACTIVATING KINASE)
DE (CAK) (39 KD PROTEIN KINASE) (P39 MO15) (STK1) (CAK1).
GN CDK7.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE; 95014737.
RA TASSAN J.-P., SCHULTZ S.J., BARTEK J., NIGG E.A.;
RL J. CELL BIOL. 127:467-478(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST;
RX MEDLINE; 94268838.
RA LEVEDAKOU E.N., HE M., BAPTIST E.W., CRAVEN R.J., CANCE W.G.,
RA WELCSH P.L., SIMMONS A., NAYLOR S.L., LEACH R.J., LEWIS T.B.,
RA LIU E.T.;
RL ONCOGENE 9:1977-1988(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95022621.
RA DARBON J.-M., DEVAULT A., TAVIAUX S., PESQUET D., MARTINEZ A.M.,
RA GALAS S., CAVADORE J.-C., DOREE M., BLANCHARD J.-M.;
RL ONCOGENE 9:3127-3138(1994).
RN [4]
RP MUTAGENESIS.
RX MEDLINE; 94349374.
RA FISHER R.P., MORGAN D.O.;
RL CELL 78:713-724(1994).
CC -1- FUNCTION: CATALYTIC PART OF CAK WHICH ACTIVATES CYCLIN-ASSOCIATED
CC CDC2/CDC2/CDK4 BY THREONINE PHOSPHORYLATION. ITS EXPRESSION AND
CC ACTIVITY ARE CONSTANT THROUGHOUT THE CELL CYCLE.
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT T-170 IS REQUIRED FOR
CC ENZYMATIC ACTIVITY (BY SIMILARITY).
CC -1- SUBUNIT: PROBABLY ASSOCIATES WITH CYCLIN H AND MAT1 TO FORM A
CC MULTIMERIC ACTIVE ENZYME (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: NUCLEAR.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN
CC KINASES. BELONGS TO THE CDC2/CDC28 SUBFAMILY.
DR EMBL; X79193; G485909; -
DR EMBL; L20320; G348243; -
DR EMBL; X77743; G468789; -
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW PHOSPHORYLATION; CELL CYCLE; CELL DIVISION; NUCLEAR PROTEIN.
FT DOMAIN 12 295
FT NP_BIND 18 26 ATP (BY SIMILARITY).
FT BINDING 41 41 ATP (BY SIMILARITY).
FT ACT_SITE 137 137 BY SIMILARITY.
FT MOD_RES 164 164 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 170 170 PHOSPHORYLATION (BY SIMILARITY).
FT MUTAGEN 41 41 K->A: TOTAL LOSS OF ACTIVITY.
FT MUTAGEN 170 170 T->A: TOTAL LOSS OF ACTIVITY.
SQ SEQUENCE 346 AA; 39038 MW; 2AA6E660 CRC32;
```

Query Match 63.5%; Score 47; DB 2; Length 346;  
Best Local Similarity 71.4%; Pred. No. 6.53e+00;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Db 301 ptpgqcl 307  
1:1:111  
QY 3 PSPACQL 9
```

RESULT 15

ID	LEW2_RAT	STANDARD;	PRT;	549 AA.
AC	P98105;			
DT	01-FEB-1996 (REL. 33, CREATED)			
DT	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)			
DT	01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)			
DE	E-SELECTIN PRECURSOR (ENDOTHELIAL LEUKOCYTE ADHESION MOLECULE 1)			
DE	(ELAM-1) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 2) (LECAM2)			
DE	(CD62E).			
GN	SELE OR ELAM-1.			
OS	RATTUS NORVEGICUS (RAT).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; RODENTIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-LUNG;			
RA	ROSENBLUM C.L., AUCHAMPACH J.A., ANDERSON D.C., MANNING A.M.;			
RL	SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.			
CC	-1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND			
CC	MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY			
CC	ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF			
CC	POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF			
CC	GLYCOLIPIDS).			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: TO OTHER SELECTINS/LECAMS.			
CC	-1- SIMILARITY: CONTAINS A C-TYPE LECTIN FAMILY DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 5 SUSHI (SCR) REPEATS.			
DR	EMBL; L25527; G409235; -			
KW	CELL ADHESION; TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN; LECTIN;			
KW	SELECTIN; SIGNAL; SUSHI; REPEAT.			
FT SIGNAL	1 21			POTENTIAL.
FT CHAIN	22 549			E-SELECTIN.
FT DOMAIN	22 494			EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	495 516			POTENTIAL.
FT DOMAIN	517 549			CYTOPLASMIC (POTENTIAL).
FT DOMAIN	38 138			C-TYPE LECTIN (SHORT FORM).
FT DOMAIN	139 175			EGF-LIKE.
FT DOMAIN	179 486			5 X SUSHI (SCR) REPEATS.
FT REPEAT	179 239			SUSHI 1.
FT REPEAT	242 301			SUSHI 2.
FT REPEAT	304 364			SUSHI 3.
FT REPEAT	367 427			SUSHI 4.
FT REPEAT	430 486			SUSHI 5.
FT DISULFID	40 138			BY SIMILARITY.
FT DISULFID	111 130			BY SIMILARITY.
FT DISULFID	143 154			BY SIMILARITY.
FT DISULFID	148 163			BY SIMILARITY.
FT DISULFID	165 174			BY SIMILARITY.
FT DISULFID	180 225			BY SIMILARITY.
FT DISULFID	210 238			BY SIMILARITY.
FT DISULFID	243 287			BY SIMILARITY.
FT DISULFID	273 300			BY SIMILARITY.
FT DISULFID	305 350			BY SIMILARITY.
FT DISULFID	336 363			BY SIMILARITY.
FT DISULFID	368 413			BY SIMILARITY.
FT DISULFID	399 426			BY SIMILARITY.
FT DISULFID	431 472			BY SIMILARITY.
FT DISULFID	458 485			BY SIMILARITY.
FT CARBOHYD	25 25			POTENTIAL.
FT CARBOHYD	60 60			POTENTIAL.
FT CARBOHYD	145 145			POTENTIAL.
FT CARBOHYD	192 192			POTENTIAL.
FT CARBOHYD	203 203			POTENTIAL.
FT CARBOHYD	266 266			POTENTIAL.
FT CARBOHYD	313 313			POTENTIAL.
FT CARBOHYD	320 320			POTENTIAL.
FT CARBOHYD	333 333			POTENTIAL.

FT CARBOHYD 441 441 POTENTIAL.  
FT CARBOHYD 465 465 POTENTIAL.  
SQ SEQUENCE 549 AA; 60079 MW; F47BC9BC CRC32;

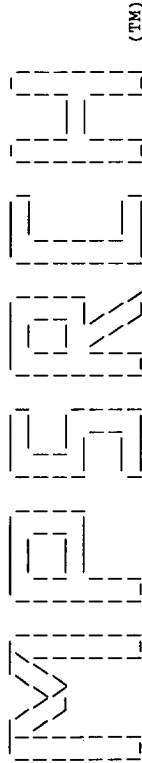
Query Match 63.5%; Score 47; DB 5; Length 549;  
Best Local Similarity 62.5%; Pred. No. 6.53e+00;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 234 papachvv 241

!::!::!  
QY 3 PSPACQLV 10

Search completed: Tue Jun 10 11:26:49 1997  
Job time : 8 secs.

\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:28:59 1997; MasPar time 1.82 Seconds  
Tabular output not generated. 59.984 Million cell updates/sec

Title: >US-08-231-565A-36  
Description: (1-10) from US08231565A.pep  
Perfect Score: 55  
Sequence: 1 SLADTNSLAV 10

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 14.540; Variance 45.010; scale 0.323

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	55	100.0	10	15	Melanoma-specific mut	5.99e+00
2	55	100.0	10	15	gp100 melanoma antige	5.99e+00
3	55	100.0	661	14	Melanoma associated a	5.99e+00
4	55	100.0	661	15	MART-1 melanoma antig	5.99e+00
5	55	100.0	661	15	MART-1 melanoma antig	5.99e+00
6	48	87.3	9	15	Melanoma-specific mut	3.59e+01
7	43	78.2	887	18	CagC antigenic polype	1.24e+02
8	40	72.7	3639	8	ACV5.	2.56e+02
9	40	72.7	3712	3	ACV synthetase.	2.56e+02
10	40	72.7	3768	3	ACV5.	2.56e+02
11	40	72.7	3778	3	ACV synthetase.	2.56e+02
12	39	70.9	256	2	PRV 28K gene product.	3.24e+02
13	39	70.9	780	9	Enzyme involved in ei	3.24e+02
14	39	70.9	3164	16	Hepatitis GB virus (H	3.24e+02
15	38	69.1	398	4	Yeast proteasome rcl	4.10e+02
16	38	69.1	398	15	Pseudomonas putida L-	4.10e+02
17	38	69.1	412	2	Tumour growth inhibit	4.10e+02
18	38	69.1	412	14	Human TGF-beta 3 prot	4.10e+02
19	38	69.1	412	4	Transforming Growth F	4.10e+02
20	38	69.1	413	4	Mutant Transforming G	4.10e+02

21	38	69.1	1422	16	R82070	Hepatitis GB virus (H	4.10e+02
22	38	69.1	2516	19	W01884	Invertebrate calcium	4.10e+02
23	38	69.1	2516	19	W01875	Neuronal invertebrate	4.10e+02
24	37	67.3	32	7	R38148	Acetobacter diguanyla	5.18e+02
25	37	67.3	43	17	R87124	Protocadherin clone M	5.18e+02
26	37	67.3	315	9	R47927	CM-PD gene product (S	5.18e+02
27	37	67.3	315	2	R24019	PAJ16 fragment.	5.18e+02
28	37	67.3	315	9	R47926	CM-PD gene product (S	5.18e+02
29	37	67.3	327	19	W03995	Murine gonadotropin-r	5.18e+02
30	37	67.3	327	19	R47164	Sequence of mouse gon	5.18e+02
31	37	67.3	345	10	R52824	GTP-cyclohydrolase II	5.18e+02
32	37	67.3	409	8	R41227	910 SLG protein.	5.18e+02
33	37	67.3	475	3	R14625	S. typhimurium HtrA p	5.18e+02
34	37	67.3	592	16	R86914	Xylanase XYNB.	5.18e+02
35	37	67.3	669	2	R23968	facA gene product.	5.18e+02
36	37	67.3	682	17	R87154	Alternatively spliced	5.18e+02
37	37	67.3	836	11	R58912	Product of alternativ	5.18e+02
38	37	67.3	857	6	R29814	S receptor kinase pro	5.18e+02
39	37	67.3	858	10	R53404	S-Locus receptor (ser	5.18e+02
40	37	67.3	1026	11	R58906	Human protocadherin-4	5.18e+02
41	37	67.3	1481	19	W03740	rchd528 gene product.	5.18e+02
42	37	67.3	1977	18	R09641	Peripheral nervous sy	5.18e+02
43	37	67.3	1984	18	R99639	Peripheral nervous sy	5.18e+02
44	37	67.3	1989	18	R92317	Peripheral nervous sy	5.18e+02
45	37	67.3	4544	9	R47861	Alpha 2-Macroglobulin	5.18e+02

ALIGNMENTS

RESULT 1  
ID R82172 standard; peptide; 10 AA.  
AC R82172;  
DT 25-MAR-1996 (first entry)  
DE Melanoma-specific mutant immunogen epitope 10mer peptide.  
KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
KW cytotoxic T cell; lymphocyte; HLA-A2.  
OS Homo sapiens.  
PN W09522561-A2.  
PD 24-AUG-1995.  
PF 16-FEB-1995; U01991.  
PR 16-FEB-1994; US-197399.  
PR 16-FEB-1994; US-234784.  
PR 29-APR-1994; US-234784.  
PA (UUVI-) UNIV VIRGINIA PATENT FOUND.  
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
DR WPI: 95-302668/39.  
PT Melanoma-specific immunogen comprises epitope(s) homologous with  
PT pmei.17 - are highly potent stimulators of HLA-A2-CTL's useful in  
PT adoptive immuno-therapy  
PS Example 8; Page 52; 148pp; English.  
CC A melanoma-specific immunogen homologous with pmei-17 comprises one  
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
CC R82108 are of particular interest. The immunogen can be used for  
CC partial protection in mammals against melanoma peptides which are  
CC homologous with pmei-17 are highly potent stimulators of HLA-A2+  
CC CTLs in several cell lines and can be used in immunotherapy or  
CC incorporated into immunogenic conjugates as vaccines.  
SQ Sequence 10 AA;

Query Match 100.0%; Score 55; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5.99e+00;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 sladtntslav 10  
|||||  
QY 1 SLADTNSLAV 10

RESULT 2  
ID R84204 standard; Peptide; 10 AA.  
AC R84204;  
DT 25-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G10-9).

KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
 KW metastatic melanoma; tumour-associated antigen;  
 KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
 KW therapy; vaccine.  
 OS Synthetic.  
 PN WO9529193-A2.

PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 55; Page 131; 184pp; English.  
 CC The immunogenic peptide is derived from CDNA25 (R84854), a  
 CC melanoma antigen derivative of gp100 (see R84855). The  
 CC peptide and its derivatives (see R84200-R84211) are used in  
 CC medicaments (vaccines) for the treatment or prevention (by  
 CC immunization) of melanoma. Antibodies against melanoma-specific  
 CC antigens and its immunogenic peptides may be used in the  
 CC detection and isolation of the antigen from a sample, the  
 CC detection of which is indicative of a disease state  
 CC (melanoma or metastatic melanoma).  
 CC Sequence 10 AA;

Query Match 100.0%; Score 55; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 5.99e+00;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 sladtntslav 10  
 Qy 1 SLADTNSLAV 10

RESULT 3

ID R78646 standard; Protein; 661 AA.  
 AC R78646;  
 DE 22-JAN-1996 (first entry)  
 DT Melanoma associated antigen gp100.  
 KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
 KW identification; tumour; gp100.  
 OS Homo sapiens.  
 PN EP-668350-A1.  
 PD 23-AUG-1995.  
 PF 14-FEB-1995; 200348.  
 PR 16-FEB-1994; EP-200337.  
 PR 21-DEC-1994; EP-203709.  
 PA (ALKU ) AKZO NOBEL NV.  
 PI Adema GJ, Figdor CG;  
 DR WPI; 95-284790/38.  
 DR N-PSDB; Q96055.  
 PT Melanoma associated antigen gp100 - used in vaccines and for the  
 PT detection of tumours  
 PS Claim 1; Page 22-24; 40pp; English.  
 CC Immunogenic peptides derived from the melanoma associated antigen  
 CC may be used in the production of vaccines. Nucleotide sequences  
 CC encoding the immunogenic peptides may be used as primers and probes  
 CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
 CC capable of binding to the melanoma associated antigen can be  
 CC cultured ex vivo and returned to melanoma particles, and when  
 CC radiolabelled, they may be used to identify tumour deposits.  
 CC Sequence 661 AA;

Query Match 100.0%; Score 55; DB 14; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 5.99e+00;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 570 sladtntslav 579  
 Qy 1 SLADTNSLAV 10

RESULT 4

ID R84855 standard; Protein; 661 AA.  
 AC R84855;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen gp100.  
 KW gp100; MART-1; melanoma antigen recognised by T-cell;  
 KW CDNA25 antigen derivative; melanocyte; melanoma;  
 KW metastatic melanoma; tumour-associated antigen; immunogen;  
 KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
 OS Mammalian sp.

Key Location/Qualifiers  
 FT Peptide 154..163  
 FT /label= G9-154\_immunogenic\_peptide  
 FT Peptide 208..217  
 FT /label= G9-209\_immunogenic\_peptide  
 FT /note= "see R84210"  
 FT Peptide 280..288  
 FT /label= G9-280\_immunogenic\_peptide  
 FT /note= "see R84208"  
 FT Peptide 457..266  
 FT /label= immunogenic\_peptide  
 FT Peptide 476..485  
 FT /label= immunogenic\_peptide  
 PN WO9529193-A2.

PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA;  
 DR WPI; 95-382963/49.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 7A; 184pp; English.  
 CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which  
 CC is recognized by T-lymphocytes, and is a derivative of the  
 CC melanoma-specific antigen CDNA25 (see R84854). gp100 is a source  
 CC of immunogenic peptides which are optionally modified to enhance  
 CC their binding to a MHC molecule, and used in medicaments,  
 CC especially vaccines, for the treatment or prevention (by  
 CC immunisation) of melanoma. Antibodies against CDNA2 and its  
 CC immunogenic peptides may be used in the detection and isolation  
 CC of the antigen from a sample, the detection of which is indicative  
 CC of a disease state (melanoma or metastatic melanoma).  
 CC Sequence 661 AA;

Query Match 100.0%; Score 55; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 5.99e+00;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 570 sladtntslav 579  
 Qy 1 SLADTNSLAV 10

RESULT 5

ID R84854 standard; Protein; 661 AA.  
 AC R84854;  
 DT 08-MAY-1996 (revised)  
 DT 20-APR-1996 (first entry)  
 DE MART-1 melanoma antigen CDNA25.  
 KW CDNA25; MART-1; melanoma antigen recognised by T-cell;  
 KW gp100 antigen derivative; melanoma; metastatic melanoma;  
 KW tumour-associated antigen; immunogen; diagnosis; prognosis;  
 KW prophylaxis; therapy; vaccine.  
 OS Mammalian sp.  
 Key Location/Qualifiers  
 FT Peptide 457..466  
 FT /label= antigenic\_peptide

FT /note= "see R84199"  
 PN WO9529193-A2.  
 PD 02-NOV-1995.  
 PF 21-APR-1995; U05063.  
 PR 22-APR-1994; US-231565.  
 PR 05-APR-1995; US-417174.  
 PA (USSH ) US SEC DEPT HEALTH.  
 PI Kawakami Y, Rosenberg SA.  
 DR WPI: 95-382963/49.  
 DR N-PSDB; T02716.  
 PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
 PT vectors, host cells and antibodies, used to detect, treat and  
 PT immunise animal against melanoma.  
 PS Claim 81; Fig 5A; 184pp; English.  
 CC cDNA2 is a melanoma antigen (MART-1) which is recognized by  
 CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-  
 CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source  
 CC of immunogenic peptides (see R84199) which are optionally modified  
 CC (see R84200-R84211) to enhance their binding to a MHC molecule and  
 CC used in medicaments, especially vaccines, for the treatment or  
 CC prevention (by immunisation) of melanoma. Antibodies against cDNA2  
 CC and its immunogenic peptides may be used in the detection and  
 CC isolation of the antigen from a sample, the detection of which is  
 CC indicative of a disease state (melanoma or metastatic melanoma).  
 SQ Sequence 661 AA;

Query Match 100.0%; Score 55; DB 15; Length 661;  
 Best Local Similarity 100.0%; Pred. No. 5.99e+00;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 570 sladtnslav 579  
 |||||  
 Qy 1 SLADTNSLAV 10

RESULT 6  
 ID R82129 standard; peptide; 9 AA.  
 AC R82129;  
 DT 25-MAR-1996 (first entry)  
 DE Melanoma-specific mutant immunogen epitope 9mer peptide.  
 KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;  
 KW cytotoxic T cell; lymphocyte; HLA-A2.  
 OS Homo sapiens.  
 PN WO9522561-A2.  
 PD 24-AUG-1995.  
 PF 16-FEB-1995; U01991.  
 PR 16-FEB-1994; US-197399.  
 PR 29-APR-1994; US-234784.  
 PA (UYVI-) UNIV VIRGINIA PATENT FOUND.  
 PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;  
 DR WPI: 95-302688/39.  
 PT Melanoma-specific immunogen comprises epitope(s) homologous with  
 PT pMEL17 - are highly potent stimulators of HLA-A2+CTL's useful in  
 PT adoptive immuno-therapy  
 PS Example 8; Page 51; 148pp; English.  
 CC A melanoma-specific immunogen homologous with pMEL-17 comprises one  
 CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-  
 CC R82194 capable of eliciting a CTL response. The epitopes R82098-  
 CC R82108 are of particular interest. The immunogen can be used for  
 CC partial protection in mammals against melanoma peptides which are  
 CC homologous with pMEL-17 are highly potent stimulators of HLA-A2+  
 CC CTLs in several cell lines and can be used in immunotherapy or  
 CC incorporated into immunogenic conjugates as vaccines.  
 SQ Sequence 9 AA;

Query Match 87.3%; Score 48; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.59e+01;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 sladtnsla 9  
 |||||  
 Qy 1 SLADTNSLA 9

RESULT 7  
 ID R95004 standard; Protein; 887 AA.  
 AC R95004;  
 DT 16-JUL-1996 (first entry)  
 DE CagC antigenic polypeptide.  
 KW CagB; CagC; antigen; infection; peptic ulcer; gastric carcinoma;  
 KW antisense oligonucleotide; treatment; prevention; diagnosis;  
 KW attenuation; vaccine.  
 OS Helicobacter pylori.  
 FH Key Location/Qualifiers  
 FT Peptide 1..22  
 FT /note= "signal peptide"  
 FT Cleavage\_site 22..23  
 FT /note= "signal peptidase cleavage site"  
 PN WO9612825-A1.  
 PD 02-MAY-1996.  
 PF 20-OCT-1995; U13659.  
 PR 21-OCT-1994; US-327494.  
 PA (UYVA-) UNIV VANDERBILT.  
 PI Blaser MJ, Sharma SA, Tummuru MKR;  
 DR WPI: 96-230623/23.  
 DR N-PSDB; T14373.  
 PT DNA encoding Helicobacter pylori CagB and CagC antigenic  
 PT polypeptide(s) - useful for diagnosing H. pylori infection and to  
 PT indicate predisposition to peptic ulceration or gastric cancer.  
 PS Claim 10; Page 59-62; 74pp; English.  
 CC The sequence represents the H. pylori CagC antigenic polypeptide.  
 CC CagC is transcribed with the CagB antigen (see R95003) as a  
 CC single transcript. Detection of these proteins using antibodies,  
 CC or detection of the DNA using specific probes or primers, can be  
 CC used to diagnose H. pylori infection and to indicate predisposition  
 CC to peptic ulcers or gastric carcinoma. The proteins and their  
 CC fragments may be used in vaccines. Mutant H. pylori deficient  
 CC in CagB (ATCC 55611) or CagC (ATCC 55612) are attenuated because  
 CC they do not stimulate synthesis of interleukin-8 in epithelial  
 CC cells.  
 SQ Sequence 887 AA;

Query Match 78.2%; Score 43; DB 18; Length 887;  
 Best Local Similarity 77.8%; Pred. No. 1.24e+02;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 715 sladstla 723  
 |||||  
 Qy 1 SLADTNSLA 9

RESULT 8  
 ID R40227 standard; Protein; 3639 AA.  
 AC R40227;  
 DT 21-FEB-1994 (first entry)  
 DE ACVS.  
 KW Delta-(L-alpha-aminoadipyl)-L-cystinyl-D-valine synthase; ACVS;  
 KW beta-lactam; antibiotic; transformed; cephalosporin; vector.  
 OS Acromonium chrysogenum.  
 PN J05192162-A.  
 PD 03-AUG-1993.  
 PF 25-JUL-1991; 186222.  
 PR 31-JUL-1990; JP-205677.  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 DR WPI: 93-277475/35.  
 DR N-PSDB; Q48231.  
 PT DNA coding delta-(L-alpha-amino-adipyl) L-cystinyl D-valine  
 PT synthase - for improved productivity of cephalosporin antibiotics  
 PS Claim 1; Page 14-27; 69pp; Japanese  
 CC The sequence (Q48231) is of a vector which includes the  
 CC delta-(L-alpha-aminoadipyl)-L-cystinyl-D-valine synthase gene.  
 CC This sequence was transformed into a host cell to express the ACVS  
 CC product. The protein produced (R40227) was then used to manufacture  
 CC a beta-lactam antibiotic.  
 SQ Sequence 3639 AA;



Query Match 72.7%; Score 40; DB 8; Length 3639;  
 Best Local Similarity 66.7%; Pred. No. 2.56e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 1382 ledtaalay 1390  
 | | | : | | |  
 QY 2 LADTNSLAV 10

## RESULT 9

ID R13896 standard; Protein; 3712 AA.  
 AC R13896; 22-NOV-1991 (first entry)  
 DE ACV synthetase.  
 KW Beta lactam antibiotics; penicillin.  
 OS Acronium chrysogenum.  
 FH Key Location/Qualifiers  
 FT Domain 301..1068  
 FT /label= I  
 FT /function= activation of amino acid substrate  
 FT Region 374..423  
 FT /label= subdomain  
 FT Region 474..501  
 FT /label= subdomain  
 FT Region 655..699  
 FT /label= subdomain  
 FT Region 725..754  
 FT /label= subdomain  
 FT Domain 1392..2154  
 FT /label= II  
 FT /function= activation of amino acid substrate  
 FT Region 1470..1518  
 FT /label= subdomain  
 FT Region 1564..1590  
 FT /label= subdomain  
 FT Region 1745..1789  
 FT /label= subdomain  
 FT Region 1817..1846  
 FT /label= subdomain  
 FT Domain 2474..3295  
 FT /label= III  
 FT /function= activation of amino acid substrate  
 FT Region 2554..2603  
 FT /label= subdomain  
 FT Region 2647..2673  
 FT /label= subdomain  
 FT Region 2827..2871  
 FT /label= subdomain  
 FT Region 2899..2928  
 FT /label= subdomain  
 FT Domain 3560..3647  
 FT /label= IV  
 FT /function= thioesterase  
 FT PN EP-445868-A.  
 PD 11-SEP-1991.  
 PF 27-FEB-1991; 200423.  
 PR 28-FEB-1990; EP-200475.  
 PR 28-FEB-1990; EP-200488.  
 PR 02-JUL-1990; EP-201768.  
 PR 03-OCT-1990; EP-202628.  
 PR 27-FEB-1991; EP-200423.  
 PA (KONN ) GIST-BROCADES NV.  
 PI Veenstra AE, Martin JF, Garcia BD, Gutierrez S, Barredo JL;  
 PI Montenegro PE, Von Doehren H, Palissa H, Van Liempt H;  
 DR WPI; 91-268735/37.  
 DR N-PSDB; Q13608.  
 PT DNA encoding amino:adipyl-cysteiny-l-valine synthetase - used for  
 prodn. of the enzyme or enhanced prodn. of new or known  
 beta-lactam antibiotic cpds.  
 PS Claim 1; Page 20; 54pp; English.  
 CC The DNA sequence was obt'd. from five subclones isolated from a  
 CC gene library of A. chrysogenum c10 (ATCC 48). The protein  
 CC sequence was deduced from the DNA. Three distinct regions of

CC homology have been identified, domains I, II and III. Within  
 CC these domains several even more conserved elements can be  
 CC distinguished. Since the enzyme synthesises a tripeptide, which  
 CC most probably requires the activation of three amino acids, a  
 CC role for these domains in the activation reactions seems likely.  
 CC A fourth domain is thought to act as a thioesterase.  
 CC The gene can be used to express the synthetase enzyme which can  
 CC be used for the prodn. of new beta-lactam antibiotics.  
 CC See also R13896.  
 SQ Sequence 3712 AA;

Query Match 72.7%; Score 40; DB 3; Length 3712;  
 Best Local Similarity 66.7%; Pred. No. 2.56e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 1455 ledtaalay 1463  
 | | | : | | |  
 QY 2 LADTNSLAV 10

## RESULT 10

ID R13753 standard; Protein; 3768 AA.  
 AC R13753; 14-NOV-1991 (first entry)  
 DE ACVS.  
 KW Delta- (L-alpha-aminoadipyl)-L-cysteiny-l-D-valine synthetase;  
 KW beta-lactam antibiotics.  
 OS Penicillium chrysogenum.  
 FH Key Location/Qualifiers  
 FT Domain 301..1068  
 FT /label= domain I  
 FT Domain 1392..2154  
 FT /label= domain II  
 FT Domain 2474..3295  
 FT /label= domain III  
 FT PN EP-444758-A.  
 PD 04-SEP-1991.  
 PF 27-FEB-1991; 200422.  
 PR 28-FEB-1990; EP-200475.  
 PR 28-FEB-1990; EP-200488.  
 PR 02-JUL-1990; EP-201768.  
 PR 03-OCT-1990; EP-202628.  
 PR 27-FEB-1991; EP-200422.  
 PA (KONN ) GIST-BROCADES NV.  
 PI Veenstra AE, Martin JF, Garcia BD, Gutierrez S, Barredo JL;  
 PI Montenegro Prieto E, Von Doehren H, Palissa H, Van Liempt H.  
 DR WPI; 91-261525/36.  
 DR N-PSDB; Q13547.  
 PT Mutant delta- (L-alpha-aminoadipyl)-L-cysteiny-l-D-valine  
 PT synthetase - used in prodn. of beta-lactam antibiotics.  
 PS Disclosure; Page 19; 56pp; English.  
 CC The amino acid sequence codes for delta- (L-alpha-aminoadipyl)-L-  
 CC cysteiny-l-D-valine synthetase (ACVS). The prods. may be used for  
 CC the enhanced expression (in vivo and in vitro) of mutant enzymes  
 CC and fermentable or known and new beta- lactam antibiotics and their  
 CC precursors, partic. antibiotics of the penam and cephem classes.  
 CC See also R13754-R13756.  
 SQ Sequence 3768 AA;

Query Match 72.7%; Score 40; DB 3; Length 3768;  
 Best Local Similarity 66.7%; Pred. No. 2.56e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 1501 ledtqalay 1509  
 | | | : | | |  
 QY 2 LADTNSLAV 10

## RESULT 11

ID R13895 standard; Protein; 3778 AA.  
 AC R13895; 22-NOV-1991 (first entry)  
 DE ACV synthetase.

KW Beta lactam antibiotics; penicillin.  
 OS Penicillin chrysoeum.  
 FH Location/Qualifiers  
 FT Domain 301..1068  
 FT /label= I  
 FT /function= activation of amino acid substrate  
 FT Region 374..423  
 FT /label= subdomain  
 FT Region 474..501  
 FT /label= subdomain  
 FT Region 655..699  
 FT /label= subdomain  
 FT Region 725..754  
 FT /label= subdomain  
 FT Domain 1392..2154  
 FT /label= II  
 FT /function= activation of amino acid substrate  
 FT Region 1470..1518  
 FT /label= subdomain  
 FT Region 1564..1590  
 FT /label= subdomain  
 FT Region 1745..1789  
 FT /label= subdomain  
 FT Region 1817..1846  
 FT /label= subdomain  
 FT Domain 2474..3295  
 FT /label= III  
 FT /function= activation of amino acid substrate  
 FT Region 2554..2603  
 FT /label= subdomain  
 FT Region 2647..2673  
 FT /label= subdomain  
 FT Region 2827..2871  
 FT /label= subdomain  
 FT Region 2899..2928  
 FT /label= subdomain  
 FT Domain 3560..3647  
 FT /label= IV  
 FT /function= thioesterase  
 FT EP-445868-A.  
 PN 11-SEP-1991. 200423.  
 PF 27-FEB-1991; EP-200475.  
 PR 28-FEB-1990; EP-200475.  
 PR 28-FEB-1990; EP-200488.  
 PR 02-JUL-1990; EP-201768.  
 PR 03-OCT-1990; EP-202628.  
 PR 27-FEB-1991; EP-200423.  
 PA (KONN ) GIST-BROCADES NV.  
 PI Veenstra AE, Martin JF, Garcia BD, Gutierrez S, Barredo JL;  
 PI Montenegro PE, Von Doehren H, Palissa H, Van Liempt H;  
 DR WPI; 91-268735/37.  
 DR N-PSDB; Q13607.  
 PT DNA encoding amino:adipyl-cysteinyI-valine synthetase - used for  
 PT prodn. of the enzyme or enhanced prodn. of new or known  
 PT beta-lactam antibiotic cpds.  
 PS Claim 1; Page 20; 54pp; English.  
 CC The isolation of DNA fragments encoding the ACV synthetase is  
 CC described in EP-357119. Cosmid HMI93 contains one such fragment  
 CC which was sequenced using the Sequenase system 2.0. The protein  
 CC sequence was deduced from the DNA. Three distinct regions of  
 CC homology have been identified, domains I, II and III. Within  
 CC these domains several even more conserved elements can be  
 CC distinguished. Since the enzyme synthesizes a tripeptide, which  
 CC most probably requires the activation of three amino acids, a  
 CC role for these domains in the activation reactions seems likely.  
 CC A fourth domain is thought to act as a thioesterase.  
 CC The gene can be used to express the synthetase enzyme which can  
 CC be used for the prodn of new beta-lactam antibiotics.  
 CC See also R13896.  
 SQ Sequence 3778 AA;

Query Match 72.7%; Score 40; DB 3; Length 3778;  
 Best Local Similarity 66.7%; Pred. No. 2.56e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 DB 1511 ledtqalay 1519  
 | | | : | | |  
 QY 2 LADTNSLAV 10  
 RESULT 12  
 ID R11419 standard; Protein; 256 AA.  
 AC R11419;  
 DT 21-MAY-1991 (first entry)  
 DE PRV 28K gene product.  
 KW Vaccine.  
 OS Pseudorabies virus strain NIA-3.  
 PN WO9102795-A.  
 PD 07-MAR-1991.  
 PF 17-AUG-1990; NL0119.  
 PR 17-AUG-1989; NL-002087.  
 PA (TEWE-) STICHT TECH WETENSC.  
 PI Dewind N, Vanzijl MM, Gielkens AL, Berns AJM;  
 DR WPI; 91-087279/12.  
 DR N-PSDB; Q11243.  
 PT Mutant pseudo-rabies virus for pigs - has protein kinase or 28 K  
 PT gene region mutation for use as vaccine.  
 PS Disclosure; Fig 3; 24pp; English.  
 CC The 28K gene does not have an essential function in the  
 CC replication of the virus and mutations can be introduced into  
 CC this region. The virus may also have a mutation in the 28K gene  
 CC region. The virus can be used to prepare vaccines for protection  
 CC of pigs against pathogens. The protein is, to a low degree,  
 CC homologous to the Us2 protein of HSV-1.  
 CC See also R11418.  
 SQ Sequence 256 AA;

Query Match 70.9%; Score 39; DB 2; Length 256;  
 Best Local Similarity 40.0%; Pred. No. 3.24e+02;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

DB 187 gltddadaldv 196  
 : | : | : | : | : |  
 QY 1 SLADTNSLAV 10  
 RESULT 13  
 ID R42454 standard; Protein; 780 AA.  
 AC R42454;  
 DT 27-MAY-1994 (first entry)  
 DE Enzyme involved in eicosapentaenoic acid (EPA) synthesis.  
 KW EPA; eicosapentaenoic acid synthetase; drug; anticoagulant;  
 KW hypolipemic; hypoglycemic; antihypertensive; anticancer; pesticide;  
 KW foodstuff; additive.  
 OS Shewanella putrefaciens.  
 PN WO9323545-A.  
 PD 25-NOV-1993.  
 PF 14-MAY-1993; J00641.  
 PR 15-MAY-1992; JP-147945.  
 PA (SAGA ) SAGAMI CHEM RES CENTRE.  
 PI Kato S, Kondo K, Yamada A, Yazawa K;  
 DR WPI; 93-386577/48.  
 DR N-PSDB; Q51128.  
 PT Gene coding for eicosa-penta:enoic acid synthetase - is isolated  
 PT from Pseudomonas, Alteromonas or Shewanella and used for  
 PT recombinant prodn. of eicosa-penta:enoic acid  
 PS Claim 8; Page 71-77; 106pp; Japanese.  
 CC EPA is useful as a drug, having anticoagulant, hypolipemic,  
 CC hypoglycemic, antihypertensive and anticancer activity. It is also  
 CC a pesticide and is useful as a nutritional foodstuff and animal feed  
 CC additive.  
 SQ Sequence 780 AA;

Query Match 70.9%; Score 39; DB 9; Length 780;  
 Best Local Similarity 66.7%; Pred. No. 3.24e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

PT proteolytic activity, used for decomposition and synthesis of  
 PS protein  
 CC Claim 1; page 1; 26pp; Japanese.  
 CC The yeast proteasome has unique proteolytic activity and can  
 CC be used in the synthesis and decomposition of protein.  
 SQ Sequence 288 AA;

Query Match 69.1%; Score 38; DB 4; Length 288;  
 Best Local Similarity 62.5%; Pred. No. 4.10e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 220 slsetnql 227  
 QY 1 SLADTNSL 8

Search completed: Tue Jun 10 11:29:09 1997  
 Job time : 10 secs.

Db 141 slactnals 149  
 QY 1 SLADTNSLA 9

RESULT 14

ID R94345 standard; Protein; 3164 AA.  
 AC R94345;  
 DT 02-JUL-1996 (first entry)  
 DE Hepatitis GB virus (HGBV) clone GB contig A protein prod.  
 DE Hepatitis GB virus (HGBV) diagnosis; treatment: vaccine;  
 KW reagents: non-A; non-B; non-C; non-D; non-E; clone; GB contig A;  
 KW tamarin; infected plasma; lambda phage; cDNA library.  
 OS Hepatitis GB virus.  
 FT Key Location/Qualifiers  
 FT Misc difference 1..3164  
 FT /note= Others correspond to degenerate or STOP  
 FT codons in T00129"  
 PN WO9521922-A2.  
 PD 17-AUG-1995.  
 PF 14-FEB-1995; U02118.  
 PR 14-FEB-1994; US-196030.  
 PR 13-MAY-1994; US-242654.  
 PR 29-JUL-1994; US-283314.  
 PR 23-NOV-1994; US-344190.  
 PR 23-NOV-1994; US-344185.  
 PR 27-JAN-1995; US-344557.  
 PA (ABBO ) ABBOTT LAB  
 PI Buljk SL, Dawson GJ, Desai SM, Erker JC, Leary TP;  
 PI Muerhoff AS, Mushahwar IK, Pilot-Matias TJ, Schlauder GG;  
 PI Simons JN;  
 DR WPI: 95-293123/38.  
 DR N-PSDB: T00129.

PT Non-A, non-B, non-C, non-D, non-E Hepatitis virus reagents - useful  
 PT for diagnosis and therapy of hepatitis GB virus  
 PS Example 9; Pages 293-307; 661pp; English.  
 CC Double stranded hepatitis GB virus (HGBV) DNA obtd. from HGBV  
 CC infected tamarin plasma, using standard procedures, was used to  
 CC prepare a lambda phage HGBV cDNA library. Clones were rescued to  
 CC from the lambda phage, searched against a sequence database and  
 CC found to be unique HGBV sequences. The clones were then used to  
 CC assemble the sequences T00129/30 (GB contig A and B) which encode  
 CC the proteins R94345-47 (the 3 possible coding strand reading  
 CC frames) and R82072, respectively. Reagents which comprise the HGBV  
 CC DNA, or its protein prods. can be used for the diagnosis, therapy  
 CC or in a vaccine to prevent HGBV infection.  
 SQ Sequence 3164 AA;

Query Match 70.9%; Score 39; DB 16; Length 3164;  
 Best Local Similarity 40.0%; Pred. No. 3.24e+02;  
 Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 2519 glsesdrav 2528  
 QY 1 SLADTNSLAV 10

RESULT 15

ID R22996 standard; Protein; 288 AA.  
 AC R22996;  
 DT 08-OCT-1992 (first entry)  
 DE Yeast proteasome YCl subunit.  
 KW Proteasome; proteolytic.  
 OS Saccaromyces cerevisiae.  
 PN J04077497-A.  
 PD 11-MAR-1992.  
 PF 19-JUL-1990; 192113.  
 PR 19-JUL-1990; JP-192113.  
 PA (SAKA ) OTSUKA PHARM KK.  
 DR WPI: 92-136766/17.  
 DR N-PSDB: Q23892.  
 PT Yeast proteasome with specific aminoacid sequence - has unique

\*\*\*\*\*  
WQSERLH  
(TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:28:30 1997; MasPar time 2.87 Seconds  
99.417 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-231-565A-36  
Description: (1-10) from US08231565A.pap  
Perfect Score: 55  
Sequence: 1 SLADTNSLAV 10

Scoring table: PAM 150  
Gap 15  
Searched: 89912 seqs, 28507787 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir50  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 21.051; Variance 24.918; scale 0.845

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	55	100.0	661	13	glycoprotein gp100 p	4.57e-02
2	55	100.0	668	13	melanocyte-specific	4.57e-02
3	52	94.5	626	14	Pmel 17 protein - mo	2.31e-01
4	47	85.5	491	14	melanoma antigen hom	3.08e+00
5	42	76.4	72	8	hypothetical protein	3.45e+01
6	42	76.4	432	10	nuclear protein - Ch	3.45e+01
7	41	74.5	165	9	ferritin like protei	5.45e+01
8	41	74.5	212	10	DNA repair protein R	5.45e+01
9	41	74.5	392	8	NASS protein - Azoto	5.45e+01
10	41	74.5	432	10	glutamate-5-semialde	5.45e+01
11	41	74.5	909	11	probable membrane pr	5.45e+01
12	41	74.5	1376	14	sodium channel prote	5.45e+01
13	41	74.5	3149	4	BPLF1 protein - huma	5.45e+01
14	40	72.7	264	5	multicatalytic endop	8.55e+01
15	40	72.7	287	10	flagellin - Listeria	8.55e+01
16	40	72.7	314	16	hemC protein - Pseud	8.55e+01
17	40	72.7	568	12	t-complex protein 1	8.55e+01
18	40	72.7	578	10	urea amidolysase - Ba	8.55e+01
19	40	72.7	1841	1	DNA-directed RNA pol	8.55e+01
20	40	72.7	1860	1	DNA-directed RNA pol	8.55e+01
21	40	72.7	3712	2	alpha-aminoadipyl-cy	8.55e+01

22	40	72.7	3746	2	YGPLV3	alpha-aminoadipyl-cy	8.55e+01
23	40	72.7	3791	2	YGPLV8	alpha-aminoadipyl-cy	8.55e+01
24	39	70.9	13	11	D56661	S-locus specific gly	1.33e+02
25	39	70.9	256	4	WMBEPN	28K protein - suid h	1.33e+02
26	39	70.9	353	11	S56262	hypothetical protein	1.33e+02
27	39	70.9	390	9	D64035	hypothetical protein	1.33e+02
28	39	70.9	426	10	S28812	hypothetical protein	1.33e+02
29	39	70.9	462	9	B57145	histidine kinase prr	1.33e+02
30	39	70.9	535	12	S44827	F54F2.2 protein - Ca	1.33e+02
31	39	70.9	550	8	S47503	uxpA protein - Pseud	1.33e+02
32	39	70.9	758	16	S62432	hypothetical protein	1.33e+02
33	39	70.9	857	2	D35905	ATP-dependent Clp pr	1.33e+02
34	39	70.9	863	9	B53303	plin porin homolog	1.33e+02
35	39	70.9	894	11	S51245	probable finger prot	1.33e+02
36	39	70.9	1681	14	A55138	sodium channel mNa2	1.33e+02
37	39	70.9	3051	12	S42373	hypothetical protein	1.33e+02
38	39	69.1	145	14	A55932	galectin-5 - rat	2.05e+02
39	38	69.1	398	8	JC4174	methionine gamma-lya	2.05e+02
40	38	69.1	451	11	S37933	aspartate transamina	2.05e+02
41	38	69.1	523	16	S61713	carboxypeptidase Y p	2.05e+02
42	38	69.1	567	6	S08329	mcf.2 protein - huma	2.05e+02
43	38	69.1	1687	16	S41742	calcium channel alph	2.05e+02
44	38	69.1	1748	1	JQ1555	RNA-directed RNA pol	2.05e+02
45	38	69.1	2464	3	QRMSPI	microtubule-associat	2.05e+02

ALIGNMENTS

RESULT 1  
ENTRY #type complete  
TITLE glycoprotein gp100 precursor, melanocyte lineage - human  
ALTERNATE\_NAMES melanoma antigen 25  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 25-May-1996  
ACCESSIONS A53668; A55753  
REFERENCE A53668  
#authors Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Fidor, C.G.  
#journal J. Biol. Chem. (1994) 269:20126-20133  
#title Molecular characterization of the melanocyte lineage-specific antigen gp100.  
#accession A53668  
#molecule\_type mRNA  
#residues 1-661 #label ADE  
REFERENCE A55753  
#authors Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Yanneli, J.R.; Adema, G.J.; Mikl, T.; Rosenberg, S.A.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1994) 91:6458-6462  
#title Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection.  
#accession A55753  
#status nucleic acid sequence not shown; not compared with conceptual translation  
#molecule\_type mRNA  
#residues 1-161, 'F', 163-661 #label RAW  
KEYWORDS glycoprotein  
SUMMARY #length 661 #molecular-weight 70255 #checksum 5487

Query Match 100.0%; Score 55; DB 13; Length 661;  
Best Local Similarity 100.0%; Pred. No. 4.57e-02;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 570 sladtntslav 579  
QY 1 SLADTNSLAV 10  
|||||||

RESULT 2  
ENTRY #type complete  
TITLE melanocyte-specific protein Pmel-17 precursor - human

```

ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS    A41234
REFERENCE      A41234
#authors      Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
              Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
              Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
              coat color locus on mouse chromosome 10 and is in a
              syntenic region on human chromosome 12.
#cross-references MUID:92021023
#accession    A41234
##status      preliminary
##molecule_type mRNA
##residues    1-668 #label KWO
##cross-references GB:M77348
SUMMARY       #length 668 #molecular-weight 70932 #checksum 6409
Query Match   100.0%; Score 55; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 4.57e-02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 570 sladtmslav 579
|||||
QY 1 SLADTNSLAV 10

RESULT 3
ENTRY      S53871      #type complete
TITLE      Pmel 17 protein - mouse
ORGANISM    #formal_name Mus musculus #common_name house mouse
DATE        27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
ACCESSIONS S53871
REFERENCE   S53871
#authors    Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
              Chintamaneni, C.; Bennett, D.; Pickard, R.T.
#journal     Nucleic Acids Res. (1995) 23:154-158
#title       Mouse silver mutation is caused by a single base insertion in
              the putative cytoplasmic domain of Pmel 17.
#accession   S53871
##status     preliminary
##molecule_type mRNA
##residues   1-626 #label KWO
SUMMARY      #length 626 #molecular-weight 65979 #checksum 5710
Query Match   94.5%; Score 52; DB 14; Length 626;
Best Local Similarity 90.0%; Pred. No. 2.31e-01;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 537 sladtmslav 546
|||||
QY 1 SLADTNSLAV 10

RESULT 4
ENTRY      A49179      #type fragment
TITLE      melanoma antigen homolog rpe1 - bovine (fragment)
ORGANISM    #formal_name Bos primigenius taurus #common_name cattle
DATE        19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
ACCESSIONS A49179
REFERENCE   A49179
#authors    Kim, R.Y.; Wistow, G.J.
#journal     Exp. Eye Res. (1992) 55:657-662
#title       The cDNA RPE1 and monoclonal antibody HMB-50 define gene
              products preferentially expressed in retinal pigment
              epithelium.
#cross-references MUID:93122163
#accession   A49179
##status     preliminary

```

```

##molecule_type nucleic acid
##residues    1-491 #label KIM
##cross-references NCBI:122438; NCBI:122439
##experimental_source retinal pigment epithelium
##note         sequence extracted from NCBI backbone
SUMMARY       #length 491 #checksum 3125
Query Match   85.5%; Score 47; DB 14; Length 491;
Best Local Similarity 80.0%; Pred. No. 3.08e+00;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 398 sladtmslav 407
|||||
QY 1 SLADTNSLAV 10

RESULT 5
ENTRY      S58489      #type fragment
TITLE      hypothetical protein 1 - Caulobacter crescentus (fragment)
ORGANISM    #formal_name Caulobacter crescentus
DATE        29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change
ACCESSIONS S58489
REFERENCE   S58489
#authors    Winzler, E.; Shapiro, L.
#journal     J. Mol. Biol. (1995) 251:346-365
#title       Use of flow cytometry to identify a Caulobacter 4.5 S RNA
              temperature-sensitive mutant defective in the cell cycle.
#accession   S58489
##status     preliminary
##molecule_type DNA
##residues    1-72 #label WIN
##cross-references EMBL:U22036
SUMMARY      #length 72 #checksum 2900
Query Match   76.4%; Score 42; DB 8; Length 72;
Best Local Similarity 77.8%; Pred. No. 3.45e+01;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 ladtmslav 9
|||||
QY 2 LADTNSLAV 10

RESULT 6
ENTRY      S20108      #type complete
TITLE      nuclear protein - Chlamydomonas reinhardtii
ORGANISM    #formal_name Chlamydomonas reinhardtii
DATE        22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
ACCESSIONS S20108
REFERENCE   S20108
#authors    Schloss, J.A.; Brown Croom, H.
#journal     J. Cell Sci. (1991) 100:877-881
#title       Normal Chlamydomonas nuclear gene structure on linkage group
              XIX.
#cross-references MUID:92268239
#accession   S20108
##status     preliminary
##molecule_type DNA
##residues    1-432 #label SCH
##cross-references EMBL:X62135
GENETICS
#introns
SUMMARY      #length 432 #molecular-weight 46862 #checksum 8554
Query Match   76.4%; Score 42; DB 10; Length 432;
Best Local Similarity 50.0%; Pred. No. 3.45e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 38 sladtmslav 47
|||||
QY 1 SLADTNSLAV 10

```

```

7
RESULT 7         #type complete
ENTRY ferritin like protein (rsgA) homolog - Haemophilus influenzae
TITLE (strain Rd KW20)
ORGANISM #formal_name Haemophilus influenzae
DATE 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
10-May-1996
ACCESSIONS B64121
REFERENCE A64000
#authors Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;
Kirkness, E.F.; Kervage, A.R.; Bult, C.J.; Tomb, J.F.;
Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.;
FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton,
M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saudek,
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann,
J.L.; Geoghagen, N.S.M.; Gnehm, C.L.; McDonald, L.A.;
Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
#journal Science (1995) 269:496-512
#title Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.
#accession B64121
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-165 #label TIGR
#cross-references GB:L42023; TIGR:HL1385
#note named as homolog to a protein from Escherichia coli
#length 165 #molecular-weight 13021 #checksum 8056
SUMMARY
Query Match 74.5%; Score 41; DB 9; Length 165;
Best Local Similarity 66.7%; Pred. NO. 5.45e+01;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Db 61 lnetslav 69
QY 2 LADTNSLAV 10

RESULT 8         #type complete
ENTRY DNA repair protein RAD51 - Methanococcus jannaschii
TITLE #formal_name Methanococcus jannaschii
ORGANISM 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
13-Sep-1996
ACCESSIONS G64331
REFERENCE A64300
#authors Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
R.D.; Sutton, G.G.; Blake, J.A.; Fitzgerald, L.M.; Clayton,
R.A.; Gocayne, J.D.; Kervage, A.R.; Dougherty, B.A.;
Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
Scott, J.L.; Geoghagen, N.S.M.; Weidman, J.F.; Fuhrmann,
J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,
J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
#journal Science (1996) 273:1058-1073
#title Complete genome sequence of the methanogenic archaeon,
Methanococcus jannaschii.
#accession G64331
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-212 #label BUL
#cross-references GB:L77117; TIGR:MU0254; CDS_PID:g1510355
GENETICS
#map_position FOR239969-240607
SUMMARY #length 212 #molecular-weight 23842 #checksum 3940

```

```

Query Match 74.5%; Score 41; DB 10; Length 212;
Best Local Similarity 77.8%; Pred. NO. 5.45e+01;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 138 laktmnlav 146
QY 2 LADTNSLAV 10

RESULT 9         #type complete
ENTRY NASS protein - Azotobacter vinelandii
TITLE #formal_name Azotobacter vinelandii
ORGANISM 07-May-1995 #sequence_revision 21-Jul-1995 #text_change
21-Jul-1995
ACCESSIONS S52248
REFERENCE S52248
#authors Gutierrez, J.C.; Ramos, F.; Ortnert, L.; Tortolero, M.
#submission submitted to the EMBL Data Library, December 1994
#description nassT, two genes involved in the induction of the
assimilatory nitrite-nitrate reductase operon (nasAB) of
Azotobacter vinelandii.
#accession S52248
#status preliminary
#molecule_type DNA
#residues 1-392 #label GUT
#cross-references EMBL:X83602
SUMMARY #length 392 #molecular-weight 42079 #checksum 883
Query Match 74.5%; Score 41; DB 8; Length 392;
Best Local Similarity 60.0%; Pred. NO. 5.45e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Db 168 pladvntlv 177
QY 1 SLADTNSLAV 10

RESULT 10        #type fragment
ENTRY glutamate-5-semialdehyde dehydrogenase (EC 1.2.1.41) -
TITLE Corynebacterium glutamicum (fragment)
ALTERNATE_NAMES gamma-glutamyl phosphate reductase
ORGANISM #formal_name Corynebacterium glutamicum
DATE 13-Jan-1995 #sequence_revision 10-Nov-1995 #text_change
10-Nov-1995
ACCESSIONS S49980
REFERENCE S49977
#authors Serebrijski, I.; Wojcik, F.; Reyes, O.; Leblon, G.
#submission submitted to the EMBL Data Library, November 1994
#description Two loci of Corynebacterium glutamicum ATCC17965 that
complement Escherichia coli mutants affected in the
expression of the proA gene product.
#accession S49980
#molecule_type DNA
#residues 1-432 #label SER
#cross-references EMBL:X82929
GENETICS
#gene proA
KEYWORDS oxidoreductase
SUMMARY #length 432 #checksum 5446
Query Match 74.5%; Score 41; DB 10; Length 432;
Best Local Similarity 50.0%; Pred. NO. 5.45e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 281 alsdsklav 290
QY 1 SLADTNSLAV 10

RESULT 11        #type complete
ENTRY S64038

```

```

TITLE      probable membrane protein YGL036w - yeast (Saccharomyces
            cerevisiae)
ALTERNATE_NAMES  hypothetical protein G3632
ORGANISM      #formal_name Saccharomyces cerevisiae
DATE          17-May-1996 #sequence_revision 17-May-1996 #text_change
            19-Jul-1996
ACCESSIONS    S64038
REFERENCE      S64003
#authors      Hebling, U.; Hofmann, B.; Delius, H.
#submission   submitted to the Protein Sequence Database, May 1996
#accession     S64038
#molecule_type DNA
#residues      1-909 #label HEB
#cross-references EMBL:272558
#experimental_source strain S288C
GENETICS
#map_position 7L
KEYWORDS      transmembrane protein
FEATURE
94-110        #domain transmembrane #status predicted #label TM1\
309-325        #domain transmembrane #status predicted #label TM2\
469-485        #domain transmembrane #status predicted #label TM3
SUMMARY       #length 909 #molecular-weight 102370 #checksum 6533

Query Match      74.5%; Score 41; DB 11; Length 909;
Best Local Similarity 70.0%; Pred. No. 5.45e+01;
Matches          7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 220 slastsslev 229
   ||| ||| |
Qy 1 SLADTNSLAV 10

RESULT 12
ENTRY      I56555 #type complete
TITLE      sodium channel protein 6 - rat
ORGANISM    #formal_name Rattus norvegicus #common_name Norway rat
DATE        26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
            26-Jul-1996
ACCESSIONS    I56555
REFERENCE      I56555
#authors      Schaller, K.L.; Krzemien, D.M.; Yarowsky, P.J.; Krueger,
            B.K.; Caldwell, J.H.
#journal      J. Neurosci. (1995) 15:3231-3242
#title        A novel, abundant sodium channel expressed in neurons and
            glia.
#cross-references MUID:95271284
#accession     I56555
#status        preliminary; translated from GB/EMBL/DDBJ
#molecule_type mRNA
#residues      1-1976 #label RES
#cross-references GB:L39018; NID:g829033; COS_PID:g829034
GENETICS
#note         gene name SCP6
SUMMARY       #length 1976 #molecular-weight 225227 #checksum 9334

Query Match      74.5%; Score 41; DB 14; Length 1976;
Best Local Similarity 60.0%; Pred. No. 5.45e+01;
Matches          6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 829 sladveglsv 838
   ||| ||| |
Qy 1 SLADTNSLAV 10

RESULT 13
ENTRY      Q088 #type complete
TITLE      BPLF1 protein - human herpesvirus 4 (strain B95-8)
ORGANISM    #formal_name human herpesvirus 4; Epstein-Barr virus
DATE        25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change
            10-Nov-1995
ACCESSIONS    G93065; A03747; S32993
REFERENCE      A93065

```

```

#authors      Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
#journal      Mol. Biol. Med. (1983) 1:21-45
#title        Sequence analysis of the 17.166 bp EcoRI fragment C of B95-8
            Epstein-Barr virus.
#cross-references MUID:85035713
#accession     G93065
#molecule_type DNA
#residues      1-3149 #label BAN
#cross-references EMBL:V01555
REFERENCE      A03794
#authors      Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.;
            Farrell, P.J.; Gibson, T.J.; Hatfull, G.; Hudson, G.S.;
            Satchwell, S.C.; Sequin, C.; Tuffnell, P.S.; Barrell, B.G.
#journal      Nature (1984) 310:207-211
#title        DNA sequence and expression of the B95-8 Epstein-Barr virus
            genome.
#cross-references MUID:84270667
#contents      annotation; protein coding region
SUMMARY       #length 3149 #molecular-weight 337955 #checksum 7480

Query Match      74.5%; Score 41; DB 4; Length 3149;
Best Local Similarity 50.0%; Pred. No. 5.45e+01;
Matches          5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 1165 pladtaalti 1174
   ||| ||| |
Qy 1 SLADTNSLAV 10

RESULT 14
ENTRY      S10318 #type complete
TITLE      multicatalytic endopeptidase complex (EC 3.4.99.46) chain
            PROS-29 - fruit fly (Drosophila melanogaster)
            proteasome chain PROS-29
ALTERNATE_NAMES #formal_name Drosophila melanogaster
ORGANISM        21-Nov-1993 #sequence_revision 26-May-1995 #text_change
DATE            03-May-1996
ACCESSIONS      S10318
REFERENCE        S10318
#authors        Haass, C.; Pesold-Hurt, B.; Kloetzel, P.M.
#journal        Nucleic Acids Res. (1990) 18:4018
#title          The Drosophila PROS-29 gene is a new member of the PROS-gene
            family.
#cross-references MUID:90326554
#accession       S10318
#status          preliminary
#molecule_type mRNA
#residues        1-264 #label HAA
#cross-references EMBL:X52319
#note           the authors translated the codon GAT for residue 151 as
            Gly
CLASSIFICATION #superfamily multicatalytic endopeptidase complex chain C9
KEYWORDS        hydrolase
SUMMARY         #length 264 #molecular-weight 29454 #checksum 4038

Query Match      72.7%; Score 40; DB 5; Length 264;
Best Local Similarity 50.0%; Pred. No. 8.55e+01;
Matches          5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 188 tiadakkdla 197
   ||| ||| |
Qy 1 SLADTNSLAV 10

RESULT 15
ENTRY      S28040 #type complete
TITLE      flagellin - Listeria monocytogenes
ORGANISM    #formal_name Listeria monocytogenes
DATE        17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change
            18-Jun-1993
ACCESSIONS    S28040
REFERENCE      S28040
#authors      Dons, L.; Rasmussen, O.F.; Olsen, J.E.

```

```
#journal      Mol. Microbiol. (1992) 6:2919-2929
#title        Cloning and characterization of a gene encoding flagellin of
               Listeria monocytogenes.
#accession    S28040
               #molecule_type DNA
               #residues      1-287 #label DON
               #cross-references EMBL:X65624
GENETICS
#gene         flaA
SUMMARY       #length 287 #molecular-weight 30445 #checksum 1009

Query Match      72.7%; Score 40; DB 10; Length 287;
Best Local Similarity 60.0%; Pred. No. 8.55e+01;
Matches          6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db    39 slddaaglav 48
      |||:||||
QY    1 SLADTNSLAV 10

Search completed: Tue Jun 10 11:28:41 1997
Job time : 11 secs.
```



\*\*\*\*\*  
M P E R E L H (TM)  
\*\*\*\*\*

\*\*\*\*\*  
Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.  
\*\*\*\*\*

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:28:02 1997; MasPar time 2.09 Seconds  
Tabular output not generated. 101.705 Million cell updates/sec

Title: >US-08-231-565A-36  
Description: (1-10) from US08231565A.pap  
Perfect Score: 55  
Sequence: 1 SLADTNSLAV 10

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 21.915; Variance 20.566; scale 1.066

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	55	100.0	668	7	PM17_HUMAN PMEL 17 PROTEIN PRECU	3.23e+03
2	41	74.5	165	4	FTN2_HAETIN FERRITIN LIKE PROTEIN	1.69e+01
3	41	74.5	432	7	PROA_CORGL GAMMA-GLUTAMYL PHOSPH	1.69e+01
4	41	74.5	638	11	YQGS_BACSU HYPOTHETICAL 73.2 KD	1.69e+01
5	41	74.5	909	11	YGD6_YEAST HYPOTHETICAL 102.4 KD	1.69e+01
6	41	74.5	3149	9	TEGU_EBV LARGE TEGUMENT PROTEIN	1.69e+01
7	40	72.7	264	7	PRC9_DRONE PROTEASOME 29 KD SUBU	2.89e+01
8	40	72.7	287	4	FLAA_LISMO FLAGELLIN.	2.89e+01
9	40	72.7	568	9	TCPO_YEAST T-COMPLEX PROTEIN 1,	2.89e+01
10	40	72.7	578	11	YCSJ_BACSU HYPOTHETICAL 63.8 KD	2.89e+01
11	40	72.7	768	5	LEM3_RAT P-SELECTIN PRECURSOR	2.89e+01
12	40	72.7	1841	8	RPLB_ARATH DNA-DIRECTED RNA POLY	2.89e+01
13	40	72.7	1860	8	RPLB_ARATH DNA-DIRECTED RNA POLY	2.89e+01
14	40	72.7	2670	11	YAQS_SCHPO PUTATIVE TRANSLATIONA	2.89e+01
15	40	72.7	3712	1	ACVS_CEPAC DELTA-(L-ALPHA-AMINO	2.89e+01
16	40	72.7	3745	1	ACVS_PENCH DELTA-(L-ALPHA-AMINO	2.89e+01
17	40	72.7	3791	1	ACVT_PENCH DELTA-(L-ALPHA-AMINO	2.89e+01
18	39	70.9	256	10	US02_PRVN3 PROTEIN US2 HOMOLOG (	4.88e+01
19	39	70.9	353	11	YF37_YEAST HYPOTHETICAL 39.9 KD	4.88e+01
20	39	70.9	390	11	YF37_YEAST HYPOTHETICAL 39.9 KD	4.88e+01
21	39	70.9	397	11	YXJA_BACSU HYPOTHETICAL 43.7 KD	4.88e+01
22	39	70.9	426	11	YL12_CORGL HYPOTHETICAL 45.7 KD	4.88e+01

23	39	70.9	481	1	ATPB_STIAU ATP SYNTHASE BETA CHA	4.88e+01
24	39	70.9	535	11	YMA2_CAEEL HYPOTHETICAL 56.5 KD	4.88e+01
25	39	70.9	552	3	DIH2_CITFR DIHYDROXYACETONE KINA	4.88e+01
26	39	70.9	758	10	YA93_SCHPO HYPOTHETICAL 85.7 KD	4.88e+01
27	39	70.9	783	2	CAD5_MOUSE VASCULAR ENDOTHELIAL-	4.88e+01
28	39	70.9	857	2	CLPB_ECOLI CLPB PROTEIN (HEAT SH	4.88e+01
29	39	70.9	865	5	HTRF_ECOLI OUTER MEMBRANE USHER	4.88e+01
30	39	70.9	1585	7	P3K3_DICDI PHOSPHATIDYLINOSITOL	4.88e+01
31	39	70.9	3051	11	YMX3_CAEEL HYPOTHETICAL 337.6 KD	4.88e+01
32	38	69.1	56	10	Y116_NPVAC HYPOTHETICAL 6.4 KD P	8.14e+01
33	38	69.1	253	3	ADH1_DROMU ALCOHOL DEHYDROGENASE	8.14e+01
34	38	69.1	256	3	FGF3_BRARE FIBROBLAST GROWTH FAC	8.14e+01
35	38	69.1	276	6	MOXJ_PARDE MOXJ PROTEIN PRECURSO	8.14e+01
36	38	69.1	287	7	PRC8_YEAST PROTEASOME COMPONENT	8.14e+01
37	38	69.1	328	4	GRHR_BOVIN GONADOTROPIN-RELEASEIN	8.14e+01
38	38	69.1	378	7	PDXB_ECOLI ERYTHRONATE-4-PHOSPHA	8.14e+01
39	38	69.1	398	6	MEGL_PSEPU METHIONINE GAMMA-LYAS	8.14e+01
40	38	69.1	451	1	AATM_YEAST ASPARTATE AMINOTRANSF	8.14e+01
41	38	69.1	523	2	CBPY_PICPA CARBOXYPEPTIDASE Y PR	8.14e+01
42	38	69.1	689	11	Y027_BPHPI HYPOTHETICAL 72.8 KD	8.14e+01
43	38	69.1	925	3	DBL_HUMAN PROTO-ONCOGENE DBL PR	8.14e+01
44	38	69.1	1538	6	LHR_ECOLI PROBABLE ATP-DEPENDEN	8.14e+01
45	38	69.1	2464	6	MAPB_MOUSE MICROTUBULE-ASSOCIATE	8.14e+01

ALIGNMENTS

RESULT	ID	PM17_HUMAN	STANDARD;	PRT;	668 AA.
AC	P40967;				
DT	01-FEB-1995	(REL. 31, CREATED)			
DT	01-FEB-1995	(REL. 31, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996	(REL. 34, LAST ANNOTATION UPDATE)			
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92021023.				
RA	KWON B.S.; CHINTAMANANI C.; KOZAK C.A.; COPELAND N.G.;				
RA	GILBERT D.J.; JENKINS N.; BARTON D.; FRANCKE U.; KOBAYASHI Y.;				
RA	KIM K.-K.;				
RL	PROC NATL. ACAD. SCI. U.S.A. 88-9228-9232(1991).				
CC	-1- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	-1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; G190106; -.				
DR	MM; 155550; -.				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23		POTENTIAL.
FT	CHAIN	24	668		PMEL 17 PROTEIN.
FT	TRANSMEM	575	595		POTENTIAL.
FT	TRANSMEM	603	623		POTENTIAL.
FT	DOMAIN	217	307		PKD.
FT	DOMAIN	315	444		10 X 13 AA TANDEM REPEATS.
FT	REPEAT	315	327		1.
FT	REPEAT	328	340		2.
FT	REPEAT	341	353		3.
FT	REPEAT	354	366		4.
FT	REPEAT	367	379		5.
FT	REPEAT	380	392		6.
FT	REPEAT	393	405		7.
FT	REPEAT	406	418		8.
FT	REPEAT	419	431		9.
FT	REPEAT	432	444		10.
FT	CARBOHYD	81	81		POTENTIAL.
FT	CARBOHYD	106	106		POTENTIAL.
FT	CARBOHYD	111	111		POTENTIAL.
FT	CARBOHYD	321	321		POTENTIAL.
FT	CARBOHYD	568	568		POTENTIAL.
SEQ	SEQUENCE	668 AA;	70992 MW;		568EIAFO CRC32;

Query Match 100.0%; Score 55; DB 7; Length 668;  
Best Local Similarity 100.0%; Pred. No. 3.23e-03;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 570 sladtmslav 579  
|||||  
QY 1 SLADTMSLAV 10

RESULT 2  
ID FTN2\_HAEIN STANDARD; PRT; 165 AA.  
AC P43708;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE FERRITIN LIKE PROTEIN 2.  
GN FTNB OR RSGA-B OR H11385.  
OS HAEMOPHILUS INFLUENZAE.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC PASTEURACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20;  
RX MEDLINE; 95350630.  
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,  
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,  
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,  
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOHAGEN N.S.M.,  
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
RA VENTER J.C.;  
RL SCIENCE 269:496-512(1995).  
CC -!- FUNCTION: IRON-STORAGE PROTEIN (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -!- SIMILARITY: VERY HIGH TO OTHER BACTERIAL EUKARYOTIC-TYPE  
CC FERRITINS, AND HIGH TO EUKARYOTIC FERRITINS.  
DR EMBL; L46018; G1007433; -.  
DR EMBL; U32819; G926461; -.  
KW IRON STORAGE; MULTIGENE FAMILY.  
FT METAL 17 17 IRON (BY SIMILARITY).  
SQ SEQUENCE 165 AA; 19021 MW; 35526D52 CRC32;

Query Match 74.5%; Score 41; DB 4; Length 165;  
Best Local Similarity 66.7%; Pred. No. 1.69e+01;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 61 lnetgslav 69  
|:|:|  
QY 2 LADTNSLAV 10

RESULT 3  
ID PROA\_CORGL STANDARD; PRT; 432 AA.  
AC P45638;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE GAMMA-GLUTAMYL PHOSPHATE REDUCTASE (GPR) (EC 1.2.1.41) (GLUTAMATE-5-  
DE SEMIALDEHYDE DEHYDROGENASE) (GLUTAMYL-GAMMA-SEMIALDEHYDE  
DE DEHYDROGENASE).  
GN PROA.  
OS CORYNEBACTERIUM GLUTAMICUM.  
OC PROKARYOTA; FIRMICUTES; IRREGULAR ASPOROGENOUS RODS; CORYNEFORM GROUP.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 17965;  
RA SEREBRIJSKI I., WOJCIK F., REYES O., LEBLON G.;  
RL SUBMITTED (NOV-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- CATALYTIC ACTIVITY: L-GLUTAMATE 5-SEMIALDEHYDE + ORTHOPHOSPHATE +  
NADP(+) = L-GAMMA-GLUTAMYL 5-PHOSPHATE + NADPH.

CC -!- PATHWAY: SECOND STEP IN PROLINE BIOSYNTHESIS PATHWAY.  
CC -!- SIMILARITY: TO OTHER PROKARYOTIC OR EUKARYOTIC GAMMA-GLUTAMYL  
CC PHOSPHATE REDUCTASES.  
DR EMBL; X82929; G599721; -.  
KW OXIDOREDUCTASE; PROLINE BIOSYNTHESIS; NADP.  
SQ SEQUENCE 432 AA; 45642 MW; 9BCCD167 CRC32;

Query Match 74.5%; Score 41; DB 7; Length 432;  
Best Local Similarity 50.0%; Pred. No. 1.69e+01;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 281 alsdsdklav 290  
:|:|:|  
QY 1 SLADTNSLAV 10

RESULT 4  
ID YQGS\_BACSU STANDARD; PRT; 638 AA.  
AC P54496;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 73.2 KD PROTEIN IN SODA-COMGA INTERGENIC REGION.  
GN YQGS.  
OS BACILLUS SUBTILIS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168 / JH642;  
RA KOBAYASHI Y., MIZUNO M., MASUDA S., TAKEMARU K., HOSONO S.,  
RA SATO T., TAKEUCHI M.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
DR EMBL; D84432; G1303866; -.  
DR SUBTILIST; BG11686; YQGS.

KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
FT TRANSMEM 5 25 POTENTIAL.  
FT TRANSMEM 42 62 POTENTIAL.  
FT TRANSMEM 71 91 POTENTIAL.  
FT TRANSMEM 154 174 POTENTIAL.  
SQ SEQUENCE 638 AA; 73207 MW; DF7D84C9 CRC32;

Query Match 74.5%; Score 41; DB 11; Length 638;  
Best Local Similarity 66.7%; Pred. No. 1.69e+01;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 213 aladnsnlt 221  
:|:|:|  
QY 1 SLADTNSLA 9

RESULT 5  
ID YGD6\_YEAST STANDARD; PRT; 909 AA.  
AC P53185;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 102.4 KD PROTEIN IN OCH1-MIG1 INTERGENIC REGION.  
GN YGL036W.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA HEBLING U., HOFMANN B., DELIUS H.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; Z72538; E243936; -.  
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.  
FT TRANSMEM 90 110 POTENTIAL.  
FT TRANSMEM 309 329 POTENTIAL.  
FT DOMAIN 377 387 POLY-ASN.  
FT DOMAIN 394 405 POLY-THR.  
FT DOMAIN 406 413 POLY-ASN.  
FT DOMAIN 756 763 POLY-GLN.

```

SQ SEQUENCE 909 AA; 102370 MW; EEB9AFD2 CRC32;
Query Match 74.5%; Score 41; DB 11; Length 909;
Best Local Similarity 70.0%; Pred. No. 1.69e+01;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 220 slatssslav 229
   ||| ||| |
QY 1 SLADTNSLAV 10

RESULT 6
ID TEGU_BV STANDARD; PRT; 3149 AA.
AC P03186;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE LARGE TEGUMENT PROTEIN.
GN BPLF1.
OS EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; GAMMAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84270667.
RA BAER R., BANKIER A.T., BIGGIN M.D., DEININGER P.L., FARRELL P.J.,
RA GIBSON T.J., HATFULL G., HUDSON G.S., SATCHWELL S.C., SEGUIN C.,
RA TUFFNELL P.S., BARRELL B.G.;
RL NATURE 310:207-211(1984).
CC -!- FUNCTION: TEGUMENT PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL36,
CC EH1-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC EMBL; V01555; G59094; -.
DR PIR; A03747; Q0B8.
DR PIR; S32993; S32993.
SQ SEQUENCE 3149 AA; 337954 MW; 8C0A1984 CRC32;

Query Match 74.5%; Score 41; DB 9; Length 3149;
Best Local Similarity 50.0%; Pred. No. 1.69e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 1165 pladtaalt1 1174
   ||| ||| |
QY 1 SLADTNSLAV 10

RESULT 7
ID PRC9_DROME STANDARD; PRT; 264 AA.
AC P18053;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE PROTEASOME 29 KD SUBUNIT (EC 3.4.99.46) (MULTICATALYTIC ENDOPEPTIDASE
DE COMPLEX 29 KD SUBUNIT).
GN PROS-29.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE; 90326554.
RA HAASS C., PESOLD-HURT B., KLOETZEL P.-M.;
RL NUCLEIC ACIDS RES. 18:4018-4018(1990).
CC -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
CC WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
CC ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
CC NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
CC PROTEOLYTIC ACTIVITY.
CC -!- PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
CC PROTEOLYTIC PATHWAY.
CC -!- SUBUNIT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL
CC SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC -!- SUBCELLULAR LOCATION: PROTEASOMES ARE FOUND IN THE CYTOPLASM AND
CC ALSO IN THE NUCLEUS.

```

---

```

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE
CC PROTEASOME A-TYPE FAMILY. BELONGS TO THE C9 SUBFAMILY.
DR EMBL; X52319; G8380; -.
DR PIR; S10318; S10318.
DR FLYBASE; FBGN0003150; PROS29.
DR PROSITE; PS00388; PROTEASOME_A.
KW PROTEASOME; HYDROLASE; PROTEASE.
SQ SEQUENCE 264 AA; 29454 MW; 2CA05F61 CRC32;

Query Match 72.7%; Score 40; DB 7; Length 264;
Best Local Similarity 50.0%; Pred. No. 2.89e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 188 tladakdlai 197
   ||| ||| |
QY 1 SLADTNSLAV 10

RESULT 8
ID FLAA_LISMO STANDARD; PRT; 287 AA.
AC Q02551;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE FLAGELLIN.
GN FLAA.
OS LISTERIA MONOCYTOGENES.
OC PROKARYOTA; FIRMICUTES; REGULAR ASPOROGENOUS ROD; UNCERTAIN.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12067;
RX MEDLINE; 93125115.
RA DONS L., RASMUSSEN O.F., OLSEN J.E.;
RL MOL. MICROBIOL. 6:2919-2929(1992).
CC -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -!- SIMILARITY: TO OTHER BACTERIAL FLAGELLINS.
DR EMBL; X65624; G44098; -.
DR PIR; S28040; S28040.
KW FLAGELLA.
SQ SEQUENCE 287 AA; 30445 MW; A8C1ADB2 CRC32;

Query Match 72.7%; Score 40; DB 4; Length 287;
Best Local Similarity 60.0%; Pred. No. 2.89e+01;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 39 siddaaglav 48
   || | |||
QY 1 SLADTNSLAV 10

RESULT 9
ID TCPQ_YEAST STANDARD; PRT; 568 AA.
AC P47079;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE T-COMPLEX PROTEIN 1, THETA SUBUNIT (TCP-1-THETA) (CCT-THETA).
GN CCT8 OR YJL008C OR J1374.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RA TO VAN D., PEREA J., JACO C.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
CC ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
CC ACTIN AND TUBULIN. IN YEAST MAY PLAY A ROLE IN MITOTIC SPINDLE
CC FORMATION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KD THAT
CC FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER MEMBERS OF TCP-1 CHAPERONIN FAMILY.

```

```

DR EMBL; Z49284; G1006721; -.
DR SGD; L0002764; CCT8.
DR PROSITE; PS00750; TCPL1.
DR PROSITE; PS00751; TCPL2.
DR PROSITE; PS00995; TCPL3.
KW CHAPERONE; ATP-BINDING; MULTIGENE FAMILY.
SQ SEQUENCE 568 AA; 61662 MW; 29618C50 CRC32;

Query Match          72.7%; Score 40; DB 9; Length 568;
Best Local Similarity 50.0%; Pred. No. 2.89e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 462 tlaetagldv 471
   :||:|:|
QY 1 SLADTNSLAV 10

RESULT 10
ID YCSJ_BAGSU          STANDARD;          PRT; 578 AA.
AC P42967;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 63.8 KD PROTEIN IN MTL0 3'REGION.
GN YCSJ.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RC STRAIN=168;
RA AKAGAWA E.; KURITA K.; SUGAWARA T.; NAKAMURA K.; YAMANE K.;
RL SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: TO H.INFLUENZAE H11730 AND H11731.
CC -!- SIMILARITY: TO YEAST UREA AMIDOLYASE (DURL1,2).
DR EMBL; D38161; G790943; -.
DR SUBTILIST; BG11231; YCSJ.
KW HYPOTHETICAL PROTEIN; ATP-BINDING.
FT NP_BIND 194 201 ATP (POTENTIAL).
SQ SEQUENCE 578 AA; 63751 MW; 987FFED CRC32;

Query Match          72.7%; Score 40; DB 11; Length 578;
Best Local Similarity 50.0%; Pred. No. 2.89e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 550 slqaeaelav 559
   || :|||
QY 1 SLADTNSLAV 10

RESULT 11
ID LEM3_RAT          STANDARD;          PRT; 768 AA.
AC P98106;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE P-SELECTIN PRECURSOR (GRANULE MEMBRANE PROTEIN 140) (GMP-140) (PADGEM)
DE (CD62P) (LEUKOCYTE-ENDOTHELIAL CELL ADHESION MOLECULE 3) (LECAM3).
GN SELP.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RX MEDLINE; 94333817.
RA AUCHAMPACH J.A.; OLIVER M.G.; ANDERSON D.C.; MANNING A.M.;
RL GENE 145:251-255(1994).
CC -!- FUNCTION: EXPRESSED AT THE SURFACE OF ACTIVATED PLATELETS AND
CC MEDIATES THE ADHESION OF NEUTROPHILS AND MONOCYTE. THE LIGAND
CC RECOGNIZED IS SIALYL-LEWIS X.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED: SPLEEN,
CC LUNG, BRAIN, LIVER, HEART, KIDNEY, THYMUS, SMALL INTESTINE.

```

---

```

CC -!- INDUCTION: ACUTE INFLAMMATION (PROBABLY).
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -!- SIMILARITY: CONTAINS A C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 8 SUSHI (SCR) REPEATS; RAT P-LECTIN LACKS THE
CC HUMAN SUSHI-2 EQUIVALENT.
DR EMBL; L23088; G349553; -.
KW CELL ADHESION; TRANSMEMBRANE; GLYCOPROTEIN; EGF-LIKE DOMAIN; LECTIN;
KW SELECTIN; SIGNAL; SUSHI; REPEAT.
FT SIGNAL 1 41 POTENTIAL.
FT CHAIN 42 768 P-SELECTIN.
FT DOMAIN 42 709 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 710 733 POTENTIAL.
FT DOMAIN 734 768 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 58 158 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 159 195 EGF-LIKE.
FT DOMAIN 199 700 8 X SUSHI (SCR) REPEATS.
FT REPEAT 261 320 SUSHI 1.
FT REPEAT 323 382 SUSHI 2.
FT REPEAT 385 444 SUSHI 3.
FT REPEAT 447 506 SUSHI 4.
FT REPEAT 509 568 SUSHI 5.
FT REPEAT 579 638 SUSHI 6.
FT REPEAT 641 700 SUSHI 7.
FT DISULFID 60 158 BY SIMILARITY.
FT DISULFID 131 150 BY SIMILARITY.
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 200 244 BY SIMILARITY.
FT DISULFID 230 257 BY SIMILARITY.
FT DISULFID 262 306 BY SIMILARITY.
FT DISULFID 292 319 BY SIMILARITY.
FT DISULFID 324 368 BY SIMILARITY.
FT DISULFID 354 381 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
FT DISULFID 416 443 BY SIMILARITY.
FT DISULFID 448 492 BY SIMILARITY.
FT DISULFID 478 505 BY SIMILARITY.
FT DISULFID 510 554 BY SIMILARITY.
FT DISULFID 540 567 BY SIMILARITY.
FT DISULFID 580 624 BY SIMILARITY.
FT DISULFID 610 637 BY SIMILARITY.
FT DISULFID 642 686 BY SIMILARITY.
FT DISULFID 672 699 BY SIMILARITY.
FT CARBOHYD 45 45 POTENTIAL.
FT CARBOHYD 54 54 POTENTIAL.
FT CARBOHYD 107 107 POTENTIAL.
FT CARBOHYD 212 212 POTENTIAL.
FT CARBOHYD 347 347 POTENTIAL.
FT CARBOHYD 456 456 POTENTIAL.
FT CARBOHYD 603 603 POTENTIAL.
FT CARBOHYD 654 654 POTENTIAL.
FT CARBOHYD 661 661 POTENTIAL.
FT CARBOHYD 679 679 POTENTIAL.
FT SITE 756 759 ENDOCYTOSIS SIGNAL (PROBABLE).
SQ SEQUENCE 768 AA; 83517 MW; CB7F68A7 CRC32;

Query Match          72.7%; Score 40; DB 5; Length 768;
Best Local Similarity 60.0%; Pred. No. 2.89e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 716 alastgslav 725
   :||:|:|
QY 1 SLADTNSLAV 10

RESULT 12
ID RPBI_ARATH          STANDARD;          PRT; 1841 AA.
AC P18616;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)

```

```

DE DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT (EC 2.7.7.6)
GN RPB205 OR RP11 OR RPB1.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPRALES; CRUCIFERAE.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. COLUMBIA;
RC MEDLINE; 9108067.
RA NAWRATH C., SCHELL J., KONCZ C.;
RL MOL. GEN. GENET. 223:65-75(1990).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
CC RNA(N).
CC -!- THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE FOUND IN
CC EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR,
CC POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S
CC AND TRNA GENES.
CC -!- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.
CC -!- THE PHOSPHORYLATION ACTIVATES POL2.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: TO OTHER RNAP II LARGE SUBUNITS.
DR EMBL; X52954; G16505; -.
DR PIR; S12071; JDMU1.
DR PROSITE; PS00115; RNA_POL_II_REPEAT.
KW DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION; ZINC; REPEAT;
KW DNA-BINDING; NUCLEAR PROTEIN; PHOSPHORYLATION; ZINC-FINGER.
FT ZN_FING 66 79 C2H2-TYPE (POTENTIAL).
FT DNA_BIND 327 398 BY SIMILARITY.
FT DOMAIN 780 790 ALPHA-AMANITIN BINDING.
FT DOMAIN 1531 1820 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
SQ SEQUENCE 1841 AA; 205002 MW; 7BAD6E95 CRC32;

Query Match 72.7%; Score 40; DB 8; Length 1841;
Best Local Similarity 75.0%; Pred. No. 2.89e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 750 slaetnsl 757
   |||:|:|
QY 1 SLADTNSL 8

RESULT 13
ID RPB0_ARATH STANDARD; PRT; 1860 AA.
AC P31635;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT (EC 2.7.7.6)
DE (VERSION 2).
GN RPB205 OR RP11 OR RPB1.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;
OC CAPRALES; CRUCIFERAE.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-CV. COLUMBIA;
RC MEDLINE; 91355869.
RA DIETRICH M.A., PRENGER J.P., GUILFOYLE T.J.;
RL PLANT MOL. BIOL. 15:207-223(1990).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
CC RNA(N).
CC -!- THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE FOUND IN
CC EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR,
CC POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S
CC AND TRNA GENES.
CC -!- PTM: THE TANDEM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.

```

```

CC THE PHOSPHORYLATION ACTIVATES POL2.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: TO OTHER RNAP II LARGE SUBUNITS.
DR EMBL; X52494; G16494; -.
DR PIR; S11960; JDMU2.
DR PROSITE; PS00115; RNA_POL_II_REPEAT.
KW DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION; ZINC; REPEAT;
KW DNA-BINDING; NUCLEAR PROTEIN; PHOSPHORYLATION; ZINC-FINGER.
FT ZN_FING 66 79 C2H2-TYPE (POTENTIAL).
FT DNA_BIND 319 390 BY SIMILARITY.
FT DOMAIN 806 816 ALPHA-AMANITIN BINDING.
FT DOMAIN 1558 1828 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.
SQ SEQUENCE 1860 AA; 207178 MW; 29A13EDF CRC32;

Query Match 72.7%; Score 40; DB 8; Length 1860;
Best Local Similarity 75.0%; Pred. No. 2.89e+01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 776 slaetnsl 783
   |||:|:|
QY 1 SLADTNSL 8

RESULT 14
ID YAO5_SCHPO STANDARD; PRT; 2670 AA.
AC Q10105;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE PUTATIVE TRANSCRIPTIONAL ACTIVATOR C18G6.05C (GCN1 HOMOLOG).
GN SPAG18G6.05C.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RP [1]
RN SEQUENCE FROM N.A.
RC STRAIN=972;
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: STRONG, TO YEAST GCN1.
DR EMBL; Z68198; E213810; -.
KW HYPOTHETICAL PROTEIN; TRANSLATION REGULATION; ACTIVATOR.
SQ SEQUENCE 2670 AA; 297333 MW; 5C214AA7 CRC32;

Query Match 72.7%; Score 40; DB 11; Length 2670;
Best Local Similarity 50.0%; Pred. No. 2.89e+01;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 1628 sltepenlav 1637
   ||:|:|:|
QY 1 SLADTNSLAV 10

RESULT 15
ID ACVS_CEPAC STANDARD; PRT; 3712 AA.
AC P25464;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE DELTA-(L-ALPHA-AMINOADIPYL)-L-CYSTEINYL-D-VALINE SYNTHETASE
DE (EC 6.-.-) (ACV SYNTHETASE) (ACVS).
GN PCBAB.
OS CEPHALOSPORIUM ACREMONIUM (ACREMONIUM CHRYSOGENUM).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; PLECTOMYCETES; EUROTIALES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91177827.
RA GUTIERREZ S., DIEZ B., MONTENEGRO E., MARTIN J.F.;
RL J. BACTERIOL. 173:2354-2365(1991).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=ATCC 11550;
RX MEDLINE; 91168300.
RA HOSKINS J.A., O'CALLAGHAN N., QUEENER S.W., CANTWELL C.A., WOOD J.S.,

```

RA CHEN V.J., SKATUD P.L.;  
RL CORR. GENET. 18:523-530(1990).  
CC -1- FUNCTION: EACH OF THE CONSTITUENT AMINO ACIDS OF THE TRIPEPTIDE  
CC ACV ARE ACTIVATED AS AMIONACYL-ADENYLATES WITH PEPTIDE BONDS  
CC FORMED THROUGH THE PARTICIPATION OF AMINO ACID THIOLESTER  
CC INTERMEDIATES.  
CC -1- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF PENICILLIN AND  
CC CEPHALOSPORIN.  
CC -1- COFACTOR: CONTAINS THREE COVALENTLY BOUND PHOSPHOPANTETHEINES  
CC (POTENTIAL).  
CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
DR PIR; A38531; YCEVC.  
DR PROSITE; PS00012; PHOSPHOPANTETHEINE.  
DR PROSITE; PS00455; AMP\_BINDING.  
KW LIGASE; ANTIBIOTIC BIOSYNTHESIS; MULTIFUNCTIONAL ENZYME;  
KW REPEAT; PHOSPHOPANTETHEINE.  
FT REPEAT 234 1062 DOMAIN 1 (ADIPATE-ACTIVATING).  
FT REPEAT 1335 2162 DOMAIN 2 (CYSTEINE-ACTIVATING).  
FT REPEAT 2409 3387 DOMAIN 3 (VALINE-ACTIVATING).  
FT BINDING 827 827 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT BINDING 1916 1916 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT BINDING 2990 2990 PHOSPHOPANTETHEINE (BY SIMILARITY).  
FT ACT\_SITE 3568 3568 THIOESTERASE (BY SIMILARITY).  
SQ SEQUENCE 3712 AA; 414767 MW; D6048F12 CRC32;

Query Match 72.7%; Score 40; DB 1; Length 3712;  
Best Local Similarity 66.7%; Pred. No. 2.89e+01;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 1455 ledtaalav 1463  
| | | : | | |  
QY 2 LADTNSLAV 10

Search completed: Tue Jun 10 11:28:12 1997  
Job time : 10 secs.

\*\*\*\*\*  
WQELH  
(TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:30:23 1997; MasPar time 1.81 Seconds  
Tabular output not generated. 54.303 Million cell updates/sec

Title: >US-08-231-565A-37  
Description: (1-9) from US08231565A.pap  
Perfect Score: 54  
Sequence: 1 SVSVSQLRA 9

Scoring table: PAM 150  
Gap 15  
Searched: 92623 seqs, 10896596 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 15.399; Variance 41.833; scale 0.368  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	54	100.0	9	15	R84205 gp100 melanoma antige	4.63e+00
2	54	100.0	661	15	R84855 MART-1 melanoma antig	4.63e+00
3	54	100.0	661	15	R84854 MART-1 melanoma antig	4.63e+00
4	54	100.0	661	14	R78646 Melanoma associated a	4.63e+00
5	43	79.6	937	17	R90539 PJG4-5-CDK-BP clone #	9.83e-01
6	42	77.8	370	3	R12965 Arg(150) CD4 mutant.	1.28e+02
7	42	77.8	466	8	R39523 Sequence of gamma-ami	1.28e+02
8	41	75.9	383	10	R51059 Sequence of plasmid p	1.67e+02
9	41	75.9	1093	9	R56978 Human myotonic dystro	1.67e+02
10	40	74.1	501	5	R27741 Sequence transcribed	2.17e+02
11	40	74.1	671	15	R85290 Streptococcus faecali	2.17e+02
12	40	74.1	3587	7	R34713 Bacillus subtilis srf	2.17e+02
13	39	72.2	14	6	R33199 Tocopherol cyclase fr	2.81e+02
14	39	72.2	250	2	P70498 Aspergillus niger tpi	2.81e+02
15	39	72.2	291	8	R39297 Maize dwarf mosaic vi	2.81e+02
16	39	72.2	339	18	R95602 STR3 (suppressor of t	2.81e+02
17	39	72.2	366	8	R41674 Rape acyl-ACP thioest	2.81e+02
18	39	72.2	706	12	R68743 BCL-6 zinc finger pro	2.81e+02
19	38	70.4	173	14	R70154 Streptococcus pneumon	3.64e+02
20	38	70.4	250	2	P70496 Aspergillus nidulans	3.64e+02

21	38	70.4	496	13	R76075 Carbamoylphosphate-sy	3.64e+02
22	38	70.4	862	15	R75787 Human wild type PMS1,	3.64e+02
23	38	70.4	862	15	R79010 Human DNA repair prot	3.64e+02
24	38	70.4	885	14	R85754 Human axl receptor (a	3.64e+02
25	38	70.4	894	14	R85753 Human axl receptor.	3.64e+02
26	38	70.4	1992	19	W04505 Moraxella 200 kba out	3.64e+02
27	37	68.5	74	13	R75497 Human C5a protein.	4.69e+02
28	37	68.5	74	2	P71666 Human anaphylatoxin C	4.69e+02
29	37	68.5	109	1	R06817 Periodontal ligament	4.69e+02
30	37	68.5	244	18	R99634 Eucalyptus SQE-1 prot	4.69e+02
31	37	68.5	280	19	W01142 TrpA protein.	4.69e+02
32	37	68.5	281	2	P70511 Brassica campestris l	4.69e+02
33	37	68.5	362	10	R54948 Brassica acyl thioest	4.69e+02
34	37	68.5	362	13	R74150 Brassica thioesterase	4.69e+02
35	37	68.5	362	6	R29172 Lactococcus lactis op	4.69e+02
36	37	68.5	383	12	R62083 Sequence of translati	4.69e+02
37	37	68.5	396	4	R20243 Adrenodoxin reductase	4.69e+02
38	37	68.5	491	2	R27787 Hepatitis C virus SK0	4.69e+02
39	37	68.5	502	13	R67591 Mouse CCT-beta subun	4.69e+02
40	37	68.5	535	16	R79578 Marek's disease virus	4.69e+02
41	37	68.5	568	13	R65496 Laminin B1k chain.	4.69e+02
42	37	68.5	1165	14	R74209 Rat IP3 receptor.	4.69e+02
43	37	68.5	2670	15	R88125 Human IP3 receptor.	4.69e+02
44	37	68.5	2670	15	R88126 Cephalosporin antibio	4.69e+02
45	37	68.5	3722	2	R10145	

ALIGNMENTS

RESULT 1  
ID R84205 standard; Peptide; 9 AA.  
AC R84205;  
DT 25-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G9-216).  
KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 55; Page 131; 184pp; English.  
CC The immunogenic peptide is derived from cDNA25 (R84854), a  
CC melanoma antigen derivative of gp100 (see R84855). The  
CC peptide and its derivatives (see R84200-R84211) are used in  
CC medicaments (vaccines) for the treatment or prevention (by  
CC immunization) of melanoma. Antibodies against melanoma-specific  
CC antigens and its immunogenic peptides may be used in the  
CC detection and isolation of the antigen from a sample, the  
CC detection of which is indicative of a disease state  
CC (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 100.0%; Score 54; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.63e+00;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 svsvsqqlra 9  
QY 1 SVSVSQLRA 9

RESULT 2  
ID R84855 standard; Protein; 661 AA.

```

AC R84855;
DT 08-MAY-1996 (revised)
DE MART-1 melanoma antigen gp100.
KW gp100; MART-1; melanoma antigen gp100.
KW cDNA25 antigen derivative; melanocyte; melanoma;
KW metastatic melanoma; tumour-associated antigen; immunogen;
KW diagnosis; prognosis; prophylaxis; therapy; vaccine.
OS Mammalian sp.
FH Key Location/Qualifiers
FT Peptide 154..163
FT /label= G9-154_immunogenic_peptide
FT Peptide 208..217
FT /label= G9-209_immunogenic_peptide
FT /note= "see R84210"
FT Peptide 280..288
FT /label= G9-280_immunogenic_peptide
FT /note= "see R84208"
FT Peptide 457..466
FT /label= immunogenic_peptide
FT Peptide 476..485
FT /label= immunogenic_peptide
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
PI WPI; 95-382963/49.
DR DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Claim 81; Fig 7A; 184pp; English.
CC gp100 is a melanocyte-melanoma-specific antigen (MART-1) which
CC is recognized by T-lymphocytes, and is a derivative of the
CC melanoma-specific antigen cDNA25 (see R84854). gp100 is a source
CC of immunogenic peptides which are optionally modified to enhance
CC their binding to a MHC molecule, and used in medicaments,
CC especially vaccines, for the treatment or prevention (by
CC immunisation) of melanoma. Antibodies against cDNA2 and its
CC immunogenic peptides may be used in the detection and isolation
CC of the antigen from a sample, the detection of which is indicative
CC of a disease state (melanoma or metastatic melanoma).
SQ Sequence 661 AA;

Query Match 100.0%; Score 54; DB 15; Length 661;
Best Local Similarity 100.0%; Pred. No. 4.63e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 216 svsvsqqlra 224
QY 1 SVSVSQLRA 9

RESULT 3
ID R84854 standard; Protein; 661 AA.
AC R84854;
DT 08-MAY-1996 (revised)
DE MART-1 melanoma antigen cDNA25.
KW cDNA25; MART-1; melanoma antigen recognised by T-cell;
KW gp100 antigen derivative; melanoma; metastatic melanoma;
KW tumour-associated antigen; immunogen; diagnosis; prognosis;
KW prophylaxis; therapy; vaccine.
OS Mammalian sp.
FH Key Location/Qualifiers
FT Peptide 457..466
FT /label= antigenic_peptide
FT /note= "see R84199"
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.

```

```

PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
DR N-PSDB; T02716.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Claim 81; Fig 5A; 184pp; English.
CC cDNA2 is a melanoma antigen (MART-1) which is recognized by
CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-
CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source
CC of immunogenic peptides (see R84199) which are optionally modified
CC (see R84200-R84211) to enhance their binding to a MHC molecule and
CC used in medicaments, especially vaccines, for the treatment or
CC prevention (by immunisation) of melanoma. Antibodies against cDNA2
CC and its immunogenic peptides may be used in the detection and
CC isolation of the antigen from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 661 AA;

Query Match 100.0%; Score 54; DB 15; Length 661;
Best Local Similarity 100.0%; Pred. No. 4.63e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 216 svsvsqqlra 224
QY 1 SVSVSQLRA 9

RESULT 4
ID R78646 standard; Protein; 661 AA.
AC R78646;
DT 22-JAN-1996 (first entry)
DE Melanoma associated antigen gp100.
KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;
KW identification; tumour; gp100.
OS Homo sapiens.
PN EP-668350-A1.
PD 23-AUG-1995.
PF 14-FEB-1995; 200348.
PR 16-FEB-1994; EP-200337.
PR 21-DEC-1994; EP-203709.
PA (ALKU ) AKZO NOBEL NV.
PI Adema GJ, Figdor CG;
PI WPI; 95-284790/38.
DR N-PSDB; Q96055.
PT Melanoma associated antigen gp100 - used in vaccines and for the
PT detection of tumours
PS Claim 1; Page 22-24; 40pp; English.
CC Immunogenic peptides derived from the melanoma associated antigen
CC may be used in the production of vaccines. Nucleotide sequences
CC encoding the immunogenic peptides may be used as primers and probes
CC in the detection of melanoma cells. Tumour infiltrating lymphocytes
CC capable of binding to the melanoma associated antigen can be
CC cultured ex vivo and returned to melanoma particles, and when
CC radiolabelled, they may be used to identify tumour deposits.
SQ Sequence 661 AA;

Query Match 100.0%; Score 54; DB 14; Length 661;
Best Local Similarity 100.0%; Pred. No. 4.63e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 216 svsvsqqlra 224
QY 1 SVSVSQLRA 9

RESULT 5
ID R90539 standard; protein; 937 AA.
AC R90539;
DT 08-AUG-1996 (first entry)

```



DE pJG4-5-CDK-BP clone #71 derived CDK4 binding protein.  
 KW Cell cycle; CDK4; regulation; G1 phase; proliferation; tumorigenesis;  
 KW cyclin dependent kinase; differentiation; CDK4 inhibitor; agonist;  
 KW antagonist.  
 OS Synthetic.  
 PN WO9533819-A2.  
 PD 14-DEC-1995.  
 PF 02-JUN-1995; U07113.  
 PR 02-JUN-1994; US-253155.  
 PR (MITO-) MITOX INC.  
 PA Braetta G, Gyuris J;  
 PI WPI; 96-040227/04.  
 DR N-PSDB; T12168.  
 PT Cyclin-dependent kinase-4 binding protein - used in the isolation of  
 PT (ant)agonists of cell cycle regulation.  
 PS Claim 1; Page 74-77; 115pp; English.  
 CC R90533-R90536 are cyclin dependent kinase 4 (CDK4) binding proteins.  
 CC encoded by clones of the plasmid pJG4-5-CDKBP. CDK4 binding proteins  
 CC (CDK4-BP) may be used in an assay for screening test compounds as  
 CC inhibitors of CDK/CDK4-BP interaction. The complexes formed by CDK4  
 CC and D-type cyclins are strongly implicated in the control of the early  
 CC G1 phase of the cell cycle and are strong candidates for controlling  
 CC and/or preventing tumorigenesis and the onset of cancer. Nucleic acids  
 CC encoding CDK4-BP or fragments of these may be used as probes/primers  
 CC to diagnose the presence or absence of genetic lesions in a gene  
 CC encoding 1 of the 24 CDK4-BP, and hence to diagnose the risk for a  
 CC subject of developing a cell-proliferation associated disorder (e.g.  
 CC cancer).  
 SQ Sequence 937 AA;

Query Match 79.6%; Score 43; DB 17; Length 937;  
 Best Local Similarity 62.5%; Pred. No. 9.83e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 483 iavsqrlr 490  
 QY ::|||::  
 2 VSVSQLRA 9

RESULT 6  
 ID R12965 standard; Protein; 370 AA.  
 AC R12965;  
 DT 18-SEP-1991 (first entry)  
 DE Arg(150) CD4 mutant.  
 KW CD4; T lymphocyte; gp120; major histocompatibility complex; AIDS;  
 KW human immunodeficiency virus.  
 PN WO9109123-A.  
 PD 27-JUN-1991.  
 PF 15-DEC-1989; U05625.  
 PR 15-DEC-1989; WO-U05625.  
 PA (DANA-) DANA FARMER CANCER.  
 PI Reinherz E, Clayton LK;  
 DR WPI; 91-208149/28.  
 PT New soluble human CD4 glyco:protein - has decreased affinity for  
 PT MHC class II antigens and is used to treat, prevent and diagnose  
 PT HIV infections  
 PS Disclosure: Table 1; 41pp; English.  
 CC Mutant M13 is one of 17 modified CD4 proteins analysed for class II  
 CC MHC binding. It corresponds to domains I and II of the human CD4  
 CC protein in which Glu at position 150 is replaced with Arg. Mutant  
 CC M13 is not one of the preferred mutants of the invention as it  
 CC retains its Class II MHC binding affinity.  
 CC See R12951-R12967.  
 SQ Sequence 370 AA;

Query Match 77.8%; Score 42; DB 3; Length 370;  
 Best Local Similarity 75.0%; Pred. No. 1.28e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 143 tlvsvqlr 150  
 QY ::|||::  
 1 SVSVSQLR 8

RESULT 7  
 ID R39523 standard; Protein; 466 AA.  
 AC R39523;  
 DT 08-FEB-1994 (first entry)  
 DE Sequence of gamma-amino butyric acid (GABA) permease encoded by  
 DE gabp gene.  
 KW gamma-amino butyric acid; GABA; permease; gabp.  
 OS Escherichia coli K12.  
 PN DE4204103-A.  
 PD 19-AUG-1993.  
 PF 12-FEB-1992; 204103.  
 PR 12-FEB-1992; DE-204103.  
 PR (FARH) HOECHST AG.  
 PI Bartsch K, Schulz A;  
 DR WPI; 93-265400/34.  
 DR N-PSDB; 046055.  
 PT New gamma-amino butyric acid permease gene - for growth  
 PT inhibition of microbial or plant cells or as selectable marker  
 PS Disclosure; Fig 1; 15pp; German.  
 CC The gene gabp is obtd. from the genome of E.coli K12 by cleavage  
 CC with BamHI and HindIII. It encodes GABA permease. Overexpression of  
 CC the gene in microbial or plant cells may be used to produce  
 CC cytotoxic effects in microorganisms, plants, plant cells or seeds.  
 CC It may also be used as a selectable marker.  
 SQ Sequence 466 AA;

Query Match 77.8%; Score 42; DB 8; Length 466;  
 Best Local Similarity 71.4%; Pred. No. 1.28e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 378 iavsqrlr 384  
 QY ::|||::  
 2 VSVSQLR 8

RESULT 8  
 ID R51059 standard; Protein; 383 AA.  
 AC R51059;  
 DT 08-OCT-1994 (first entry)  
 DE Sequence of plasmid PMVO2.  
 KW Plasmid PMVO2; Lactococcal plasmid; theta-type plasmid;  
 KW food grade plasmid.  
 OS Lactococcus lactis subspecies cremoris Wg2.  
 PN WO9406917-A.  
 PD 31-MAR-1994.  
 PF 20-SEP-1993; E02558.  
 PR 18-SEP-1992; EP-202869.  
 PA (UNIL) UNILEVER NV.  
 PA (UNIL) UNILEVER PLC.  
 PI Bron S, Kiewiet R, Seegers JFML, Venema G;  
 DR WPI; 94-118466/14.  
 DR N-PSDB; 045288.  
 PT Protein prodn. from lactic acid bacteria - transformed with food  
 PT grade plasmid having theta replication mechanism and stably  
 PT maintained under non-selective conditions  
 PS Example; Page 29-33; 94pp; English.  
 CC The inventors claim a food grade plasmid in which the replicon and  
 CC any DNA sequences required for stable maintenance in the bacterium  
 CC are derived from PMVO2 originating from Lactococcus lactis subsp.  
 CC Wg2. PMVO2 is a cryptic plasmid of 3.8kb. Its replication via a  
 CC theta-mechanism. Possible promoter and RBS sites upstream of the ORF  
 CC are given in Ft and the direct repeats in from of the rep protein  
 CC are indicated. The ORF has extensive similarity to the ORF encoding  
 CC the replication gene of pCI305 and a number of other Lactococcal  
 CC plasmids. Analogous to pCI305 the ORF was denoted ORF repB. Like  
 CC pCI305, the replication gene is preceded by a 22 bp 3 1/2 direct  
 CC repeat upstream of which an Ar-rich sequence is present. Outside  
 CC this region, a 1.8kb stretch was found that showed three different  
 CC repetitive sequences that varied in length from 50 - 200 bp.  
 CC Preceding and just overlapping the first repeated sequence of 50 bp  
 CC a 13 bp sequence is found that also appears in both the 50 bp  
 CC repeat and in the 100 bp repeat, but is absent from the 200 bp

```

CC repeat.
SQ Sequence 383 AA;
  Query Match 75.9%; Score 41; DB 10; Length 383;
  Best Local Similarity 55.6%; Pred. No. 1.67e+02;
  Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 199 siskelrt 207
  :||| ||:
QY 1 SVSVSQLRA 9

RESULT 9
ID R56978 standard; Protein; 1093 AA.
AC R56978;
DT 25-FEB-1994 (first entry)
DE Human myotonic dystrophy gene protein.
KW Abnormality; muscular dystrophy; CHR 19; chromosome 19;
KW protein kinase; polymerase chain reaction; brain.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Region 1..1093
FT /note= "encoded by predicted reading frame b,
FT x's in the sequence indicate stop codons
FT in the reading frame"
PN W09317104-A.
PD 02-SEP-1993.
PF 19-FEB-1993; U01545.
PR 20-FEB-1992; US-839255.
PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.
PI Brook JD, Housman DE.
DR WPI; 93-288410/36.
PT DNA sequence of myotonic dystrophy gene - used to produce probes
PT and identify CHR 19 abnormality and protein kinase responsible
PS Disclosure; Fig 6; 64pp; English.
CC The sequence is that encoded by predicted reading frame b of
CC the human myotonic dystrophy (DM) gene. It may be used in the
CC identification of individuals affected by DM.
SQ Sequence 1093 AA;
  Query Match 75.9%; Score 41; DB 9; Length 1093;
  Best Local Similarity 62.5%; Pred. No. 1.67e+02;
  Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 104 gvtvxqlr 111
  :||| ||:
QY 1 SVSVSQLR 8

RESULT 10
ID R27741 standard; Protein; 501 AA.
AC R27741;
DT 09-MAR-1993 (first entry)
DE Sequence transcribed from first reading frame of
DE Vaccinia virus DNA from positions 23501-25000.
KW virus vector; vaccinia virus; papillomavirus; HPV;
KW immunotherapeutic; neutral site.
OS Vaccinia virus.
PN W09216636-A.
PD 01-OCT-1992.
PF 10-MAR-1992; G00424.
PR 14-MAR-1991; GB-005383.
PA (IMMU ) IMMUNOLOGY LTD.
PI Boursnell MEG, Inglis SC, Munro AJ;
DR WPI; 92-349219/42.
DR N-PSDB; Q29469.
PT Recombinant virus vectors encoding human papillomavirus proteins
PT - for treating and vaccinating against HPV infections and
PT conditions caused by them, such as cervical cancer
PS Disclosure; Fig 21; 83pp; English.
CC To make a recombinant virus vector comprising human papillomavirus
CC genes inserted into the vaccinia virus genome, neutral sites
CC for insertion must be utilised such that replicative ability is not
adversely affected. The neutral sites are identified by analysing
the viral genome to identify ORFs which are likely to encode
functional genes and selecting sites between such ORFs or within
sequences for non-functional genes. The sequence shown is that
transcribed from the vaccinia virus WR strain positions 23501-25000
contg. the regions covered by the four fragments Salf, G, H and I.
The sequence was transcribed in all three reading frames to determine
genuine vaccinia virus genes via codon usage, thus determining neutral
sites. HPV DNA sequences may be inserted neutral sites, e.g. those
encoding E6 or E7 of HPV 16 and 18 or mutants of these proteins.
CC The recombinant virus vector may be used immunotherapeutically to
CC activate cells of the immune system against HPV.
CC See also R27723-43.
SQ Sequence 501 AA;
  Query Match 74.1%; Score 40; DB 5; Length 501;
  Best Local Similarity 66.7%; Pred. No. 2.17e+02;
  Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 144 svsakqlrt 152
  ||| ||:
QY 1 SVSVSQLRA 9

RESULT 11
ID R85290 standard; Protein; 671 AA.
AC R85290;
DT 02-APR-1996 (first entry)
DE Streptococcus faecalis autolysin.
KW Lysin; autolysin; culture; lactic acid bacteria; fermentation;
KW cheese; foodstuffs; induction.
OS Streptococcus faecalis.
PN W09531361-AI.
PD 23-NOV-1995.
PF 12-MAY-1995; N0170.
PR 12-MAY-1994; EP-201353.
PA (UNIL ) QUEST INT BV.
PI Buist G, Kok J, Ledebroer AM, Venema G;
DR WPI; 96-010946/01.
PT Lysis of a culture of lactic acid bacteria in, e.g. cheese
PT production - by in situ prodn. of an auto:lysin, regulated by an
PT inducible promoter.
PS Disclosure; Page 66-69; 103pp; English.
CC In situ production of a homologous autolysin or a heterologous
CC autolysin from a food grade Gram positive bacteria, can be used in
CC a process for the lysis of a culture of lactic acid bacteria. The
CC process can be used in the manufacture of products containing
CC cultures of lactic acid bacteria e.g. cheese, where the culture is
CC lysed following the completion of fermentation. The enhanced
CC induction of the autolysin is performed some hours after the
CC fermentation is finished. No extra lysin needs to be added and the
CC lysin does not need to be isolated or encapsulated. The time of
CC lysis can be precisely controlled. This is the Streptococcus
CC faecalis autolysin.
SQ Sequence 671 AA;
  Query Match 74.1%; Score 40; DB 15; Length 671;
  Best Local Similarity 55.6%; Pred. No. 2.17e+02;
  Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 379 gvsvanlrs 387
  :||| ||:
QY 1 SVSVSQLRA 9

RESULT 12
ID R34713 standard; Protein; 3587 AA.
AC R34713;
DT 17-AUG-1993 (first entry)
DE Bacillus subtilis srfa operon ORF2 prod.
KW Mulienzyme complex; surfactin synthetase; MCSS; ORF; surfactant.
OS Bacillus subtilis.
FH Key Location/Qualifiers

```

PT Peptide 143..149  
 FT /note= "repeat sequence"  
 FT Peptide 613..623  
 FT /note= "repeat sequence"  
 FT Peptide 855..874  
 FT /note= "repeat sequence"  
 FT Peptide 1182..1188  
 FT /note= "repeat sequence"  
 FT Peptide 1644..1654  
 FT /note= "repeat sequence"  
 FT Peptide 1900..1919  
 FT /note= "repeat sequence"  
 FT Peptide 2225..2231  
 FT /note= "repeat sequence"  
 FT Peptide 2697..2708  
 FT /note= "repeat sequence"  
 FT Peptide 2929..2948  
 FT /note= "repeat sequence"  
 FT Peptide 3255..2361  
 FT /note= "repeat sequence"  
 PN EP-540074-A.  
 PD 05-MAY-1993.  
 PF 03-OCT-1992; 203037.  
 PR 09-OCT-1991; IT-MI2683.  
 PR 02-SEP-1992; IT-MI2044.  
 PA (ENIE ) ENIRICERCHE SPA.  
 PI Carrera P, Cosmina P, De FERRA F, Grandi G, Perego M;  
 PI Rodriguez F;  
 PI WPI: 93-145447/18.  
 DR N-PSDB: Q40706.  
 PT Multi-enzyme complex surfactin synthetase DNA - is isolated from  
 PT Bacillus subtilis, and used for prodn. of surfactin for use in  
 PT pharmaceutical prods.  
 PS Claim 15; Page 35-42; 70pp; English.  
 CC The Bacillus subtilis chromosomal DNA region comprises the srfA  
 CC operon which encodes the multienzyme complex surfactin synthetase  
 CC (MCSS). Analysis of the sequence showed four regions potentially  
 CC coding for proteins, a zone upstream of the first ORF contg. the  
 CC srfA operon promoter, and a presumed terminator positioned downstream  
 CC of the stop codon of the fourth ORF. ORF2 encodes a protein  
 CC (shown) which can be divided into 3 adjacent repeat regions with  
 CC internal homology followed by a region (module B) of 500 amino acids  
 CC very similar to that of ORF1 and homologous to tyrocidin and gramicidin  
 CC synthetase subunits I.  
 CC See also R34712-21.  
 CC Sequence 3587 AA;  
 SQ

Query Match 74.1%; Score 40; DB 7; Length 3587;  
 Best Local Similarity 71.4%; Pred. No. 2.17e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 910 vavselr 916  
 :||:|  
 QY 2 VSVSQLR 8

RESULT 13  
 ID R33199 standard; peptide; 14 AA.  
 AC R33199;  
 DT 05-JUL-1993 (first entry)  
 DE Tocopherol cyclase fragment.  
 KW TC; phytol benzoquinol; enantiomer selective; green algae; wheat;  
 KW leaves.  
 OS Chlorella protothecoides.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 1  
 FT /note= "unknown"  
 FT Misc\_difference 2  
 FT /note= "unknown"  
 PN EP-531639-A.  
 PD 17-MAR-1993.  
 PF 26-JUN-1992; 110874.  
 PR 18-JUL-1991; EP-112006.

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 PI Gruening F, Hochuli E, Matzinger PK;  
 DR WPI: 93-086769/11.  
 PT New tocopherol cyclase in homogeneous form - which acts on phytol  
 PT benzo:quinol derivs. to produce R',R',R'-tocopherol cpds.  
 PT enantioselectively  
 PS Disclosure: Page 10; 22pp; English.  
 CC The novel tocopherol cyclase enzyme may be isolated from green algae,  
 CC e.g. Dunaliella salina or Chlorella protothecoides or wheat leaves  
 CC by cultivating these cells, disrupting them and isolating and  
 CC purifying TC from the soluble cellular fraction. TC has a mol. wt.  
 CC of 48-50 kD, an opt. pH range of 6.5-7, opt. temp. of 25-35 deg. C,  
 CC a stabilisation in glycerol of 8-12 percent and contains the peptide  
 CC fragment shown (some residues uncertain). The TC can be used to  
 CC convert phytol benzoquinol derivs. to produce R',R',R'-tocopherols  
 CC enantioselectively.  
 CC See also R33193-203.  
 SQ Sequence 14 AA;

Query Match 72.2%; Score 39; DB 6; Length 14;  
 Best Local Similarity 55.6%; Pred. No. 2.81e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 3 avyvaqlrg 11  
 :||:|  
 QY 1 SVSVSQLRA 9

RESULT 14  
 ID P70498 standard; protein; 250 AA.  
 AC P70498;  
 DT 27-FEB-1991 (first entry)  
 DE Aspergillus niger tpiA gene.  
 KW tpiA: triose-phosphate-isomerase; enzyme; promoter.  
 OS Aspergillus niger.  
 PN WO8704464-A.  
 PD 30-JUL-1987.  
 PR 15-JAN-1986; US-820519.  
 PR 17-JAN-1986; US-946873.  
 PR 13-JAN-1987; US-946873.  
 PA (ZYMO-) ZYMOGENETICS INC.  
 PI McKnight GL, Upshall A;  
 DR WPI: 87-221265/31.  
 DR N-PSDB: N70815.  
 PT New recombinant plasmids to integrate in Aspergillus chromosomal  
 PT DNA - useful for expressing higher eucaryotic genes esp. when  
 PT plasmid contains transcriptional promoter, eg DNA encoding  
 PT alcohol dehydrogenase  
 PS Di Di Disclosure; Fig 8; 79pp; English.  
 CC This sequence is used for expressing higher eucaryotic genes in  
 CC Aspergillus through the use of a recombinant plasmid capable of  
 CC integration into the chromosome DNA of Aspergillus. See also  
 CC N70813-14, N70816 and P70496-97.  
 CC Sequence 250 AA;  
 SQ

Query Match 72.2%; Score 39; DB 2; Length 250;  
 Best Local Similarity 71.4%; Pred. No. 2.81e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 77 isveqlr 83  
 :||:|  
 QY 2 VSVSQLR 8

RESULT 15  
 ID R39297 standard; Protein; 291 AA.  
 AC R39297;  
 DT 12-JAN-1994 (first entry)  
 DE Maize dwarf mosaic virus strain A (MDMV-A) coat protein.  
 KW MDMV; coat protein; potyvirus; Zea mays; ss.  
 OS Maize dwarf mosaic virus.  
 PN WO9314210-A.  
 PD 22-JUL-1993.

```

PF 28-DEC-1992; EP03001.
PR 08-JAN-1992; US-817922.
PA (SANO ) SANDOZ LTD.
PA (SANO ) SANDOZ PATENT GMBH.
PA (SANO ) SANDOZ-ERFINDUNGEN VERW GES MBH.
PI Clark JM, Jilka JM, Murry LE, Scarafia LEC;
DR WPI; 93-243225/30.
DR N-PSDB; Q46775.
PT Transgenic corn plant cells - comprise e.g. coat protein gene of
PT maize dwarf mosaic viral strain, produces plant resistant to
PT viral challenge
PT Disclosure; Page 29-30; 69pp; English.
PS The maize dwarf mosaic virus coat protein gene is used to transform
CC corn plant cells. These are then used to produce plants resistant
CC to a variety of viruses, particularly potyviruses and especially
CC maize dwarf mosaic virus (MDMV). Other viral sequences which may be
CC used with similar effect are antisense sequences (transcription
CC inhibitors), helicases, defective interfering sequences, intergenic
CC regions, 3' untranslated terminal sequences, viral helper components,
CC viral movement protein genes, viral satellite genes, stem-loop
CC sequences and transcriptional or translational factors of viral
CC origin.
SQ Sequence 291 AA;

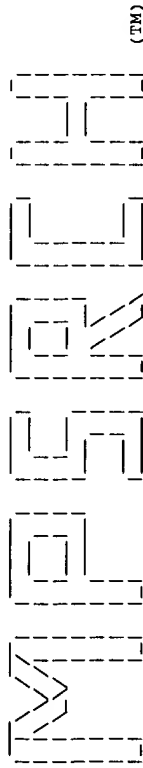
Query Match 72.2%; Score 39; DB 8; Length 291;
Best Local Similarity 66.7%; Pred. No. 2.81e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 64 svsvpkika 72
QY 1 SVSVSQLRA 9
||||:|:|

```

Search completed: Tue Jun 10 11:30:33 1997  
 Job time : 10 secs.

\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:29:54 1997; MasPar time 2.61 Seconds

Tabular output not generated.  
98.313 Million cell updates/sec

Title: >US-08-231-565A-37  
Description: (1-9) from US08231565A.pep  
Perfect Score: 54  
Sequence: 1 SVSVSQLRA 9

Scoring table: PAM 150  
Gap 15

Searched: 89912 seqs, 28507787 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p1r50

1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 21.334; Variance 25.441; scale 0.839

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	54	100.0	661	13	A53668	glycoprotein gp100 p
2	54	100.0	668	13	A41234	melanocyte-specific
3	46	85.2	491	14	A49179	melanoma antigen hom
4	46	85.2	626	14	S53871	Pmel 17 protein - mo
5	43	79.6	256	5	A32296	ubiquinol--cytochrom
6	43	79.6	274	1	A34660	ubiquinol--cytochrom
7	43	79.6	937	13	S42366	Ion proteinase homol
8	43	79.6	962	13	S57342	ATP-dependent protea
9	42	77.8	466	9	S47017	gamma-aminobutyrate
10	42	77.8	1165	12	S58236	pyruvate oxidoreduct
11	42	77.8	1729	11	S57596	hypothetical protein
12	42	77.8	2763	4	WZBE22	gene 22 protein - hu
13	42	77.8	4140	14	A39638	plectin - rat
14	41	75.9	111	10	D49692	hypothetical protein
15	41	75.9	216	11	S61701	hypothetical protein
16	41	75.9	328	2	DWHUT	L-serine dehydratase
17	41	75.9	383	9	S39651	repB protein - Lacto
18	41	75.9	753	11	S48261	hypothetical protein
19	41	75.9	958	4	P1BVCC	1a protein - cowpea
20	41	75.9	961	4	P1BVA	1a protein - brome m
21	41	75.9	4450	6	JX0340	gramicidin S synthas

22	41	75.9	4452	2	YGBSG2	gramicidin S synthet	6.76e+01
23	40	74.1	93	6	C33542	islet amyloid polype	1.05e+02
24	40	74.1	93	2	TCRTIA	islet amyloid polype	1.05e+02
25	40	74.1	148	10	D64386	ribosomal protein S1	1.05e+02
26	40	74.1	165	13	A28138	interferon-induced 1	1.05e+02
27	40	74.1	167	8	JQ1797	B3R protein - vaccin	1.05e+02
28	40	74.1	240	12	A46587	hemoglobin (linker c	1.05e+02
29	40	74.1	671	9	A38109	hemoglobin - Enteroc	1.05e+02
30	40	74.1	745	11	S57048	hypothetical protein	1.05e+02
31	40	74.1	891	8	H36790	hypothetical protein	1.05e+02
32	40	74.1	1171	1	QOKBFP	pyruvate (flavodoxin	1.05e+02
33	40	74.1	1197	9	I39613	nitrogen fixation pr	1.05e+02
34	40	74.1	1325	16	S16129	dynamin-associated pr	1.05e+02
35	40	74.1	3072	10	S35518	surfactin synthetase	1.05e+02
36	40	74.1	3583	10	S46968	surfactin synthetase	1.05e+02
37	40	74.1	3587	10	I40486	srfA2 protein - Bac	1.05e+02
38	39	72.2	255	6	S45563	3-dehydroquinone deh	1.62e+02
39	39	72.2	480	5	A42383	RNA-directed DNA pol	1.62e+02
40	39	72.2	484	10	E64432	spore coat polysacch	1.62e+02
41	39	72.2	738	10	I40719	isocitrate dehydroge	1.62e+02
42	39	72.2	764	2	QRHUGS	secretory component	1.62e+02
43	39	72.2	892	12	A41697	nitrate assimilation	1.62e+02
44	39	72.2	944	12	S48821	probable membrane pr	1.62e+02
45	39	72.2	1313	7	A48467	myosin heavy chain -	1.62e+02

ALIGNMENTS

RESULT	1	A53668	#type complete
ENTRY		glycoprotein gp100 precursor, melanocyte lineage - human	
TITLE		melanoma antigen 25	
ALTERNATE_NAMES		#formal_name Homo sapiens #common_name man	
ORGANISM		07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change	
DATE		25-May-1996	
ACCESSIONS		A53668; A55753	
REFERENCE		A53668	
#authors		Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G.	
#journal		J. Biol. Chem. (1994) 269:20126-20133	
#title		Molecular characterization of the melanocyte lineage-specific antigen gp100.	
#accession		A53668	
##molecule_type		mRNA	
##residues		1-661	##label ADE
REFERENCE		A55753	
#authors		Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Yannelis, J.R.; Adema, G.J.; Miki, T.; Rosenberg, S.A.	
#journal		Proc. Natl. Acad. Sci. U.S.A. (1994) 91:6458-6462	
#title		Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection.	
#accession		A55753	
##status		nucleic acid sequence not shown; not compared with conceptual translation	
##molecule_type		mRNA	
##residues		1-161, 'F', 163-661	##label KAW
KEYWORDS		glycoprotein	
SUMMARY		#length 661 #molecular-weight 70255 #checksum 5487	
Query Match		100.0%; Score 54; DB 13; Length 661;	
Best Local Similarity		100.0%; Pred. No. 1.09e+01;	
Matches		9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db		216 svsvsqlra 224	
QY		1 SVSVSQLRA 9	
RESULT	2	A41234	#type complete
ENTRY		melanocyte-specific protein Pmel-17 precursor - human	
TITLE			

```

ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS    A41234
REFERENCE      A41234
#authors      Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
               Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
               Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
               coat color locus on mouse chromosome 10 and is in a
               syntenic region on human chromosome 12.
#cross-references MUID:92021023
#accession    A41234
##status      preliminary
##molecule_type mRNA
##residues    #length 668 #molecular-weight 70932 #checksum 6409
#cross-references GB:M77348
SUMMARY
Query Match      100.0%; Score 54; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 1.09e-01;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 216 svsvsqr 224
Qy 1 SVSVSQLRA 9

RESULT 3
ENTRY   #type fragment
TITLE   melanoma antigen homolog rpe1 - bovine (fragment)
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE     19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
23-Mar-1995
ACCESSIONS A49179
REFERENCE   A49179
#authors    Kim, R.X.; Wistow, G.J.
#journal    Exp. Eye Res. (1992) 55:657-662
#title      The cDNA RPE1 and monoclonal antibody HMB-50 define gene
               products preferentially expressed in retinal pigment
               epithelium.
#cross-references MUID:93122163
#accession  A49179
##status    preliminary
##molecule_type nucleic acid
##residues  1-491 #label KIM
##cross-references NCBIN:122438; NCBIP:122439
##experimental_source retinal pigment epithelium
##note      #length 491 #checksum 3125
SUMMARY
Query Match      85.2%; Score 46; DB 14; Length 491;
Best Local Similarity 88.9%; Pred. No. 6.55e-00;
Matches          8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 59 svsvsqr 67
Qy 1 SVSVSQLRA 9

RESULT 4
ENTRY   #type complete
TITLE   Pmel 17 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
03-Nov-1995
ACCESSIONS S53871
REFERENCE   S53871
#authors    Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
               Chintamaneni, C.; Bennett, D.; Pickard, R.T.
#journal    Nucleic Acids Res. (1995) 23:154-158
#title      Mouse silver mutation is caused by a single base insertion in

```

```

#accession    S53871
##status      preliminary
##molecule_type mRNA
##residues    #length 626 #molecular-weight 65979 #checksum 5710
SUMMARY
Query Match      85.2%; Score 46; DB 14; Length 626;
Best Local Similarity 88.9%; Pred. No. 6.55e-00;
Matches          8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 216 svsvsqr 224
Qy 1 SVSVSQLRA 9

RESULT 5
ENTRY   #type fragment
TITLE   ubiquinol--cytochrome-c reductase (EC 1.10.2.2) iron-sulfur
               protein precursor - rat (fragment)
ALTERNATE_NAMES Rheske iron-sulfur protein
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE     17-Aug-1989 #sequence_revision 17-Aug-1989 #text_change
01-Mar-1996
ACCESSIONS A32296
REFERENCE   A32296
#authors    Nishikimi, M.; Hosokawa, Y.; Toda, H.; Suzuki, H.; Ozawa, T.
#journal    Biochem. Biophys. Res. Commun. (1989) 159:19-25
#title      Cloning and sequence analysis of a cDNA encoding the Rheske
               iron-sulfur protein of rat mitochondrial cytochrome bcl
               complex.
#cross-references MUID:89165853
#accession  A32296
##status      preliminary
##molecule_type mRNA
##residues    1-256 #label NIS
##cross-references GB:M24542
CLASSIFICATION #superfamily ubiquinol--cytochrome-c reductase iron-sulfur
               protein
KEYWORDS      2Fe-2S; electron transfer; iron-sulfur protein; membrane
               protein; oxidoreductase
FEATURE
199,204,218,220 #binding_site 2Fe-2S cluster (Cys) (covalent) #status
               predicted
SUMMARY      #length 256 #checksum 4400
Query Match      79.6%; Score 43; DB 5; Length 256;
Best Local Similarity 75.0%; Pred. No. 2.72e-01;
Matches          6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 171 avevsqr 178
Qy 1 SVSVSQLR 8

RESULT 6
ENTRY   #type complete
TITLE   ubiquinol--cytochrome-c reductase (EC 1.10.2.2) iron-sulfur
               protein precursor - bovine
ALTERNATE_NAMES cytochrome bcl complex chain 9; cytochrome bcl complex
               iron-sulfur protein; Rheske iron-sulfur protein
CONTAINS        ubiquinol--cytochrome-c reductase 8K protein;
               ubiquinol--cytochrome-c reductase iron-sulfur protein
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE     31-Mar-1991 #sequence_revision 22-Apr-1995 #text_change
05-Jan-1996
ACCESSIONS A46063; A34660; A24011; S00003; S14162
REFERENCE   A46063
#authors    Brandt, U.; Yu, L.; Yu, C.A.; Trumpower, B.L.
#journal    J. Biol. Chem. (1993) 268:8387-8390
#title      The mitochondrial targeting presequence of the Rheske
               iron-sulfur protein is processed in a single step after
               insertion into the cytochrome bcl complex in mammals and

```

```

    retained as a subunit in the complex.
#cross-references MUID:93231976
#accession A45063
##molecule_type mRNA
##residues 1-274 ##label BRA
##cross-references NCBI:129525; NCBI:129529
##experimental_source heart
#note
sequence extracted from NCBI backbone
REFERENCE
A34660
#authors
Usui, S.; Yu, L.; Yu, C.A.
#journal
Biochem. Biophys. Res. Commun. (1990) 167:575-579
#title
Cloning and sequencing of a cDNA encoding the Rieske
iron-sulfur protein of bovine heart mitochondrial
ubiquinol-cytochrome c reductase.
#cross-references MUID:90211231
#accession A34660
##molecule_type mRNA
##residues 1-8,'RHSR',13,'SYRPPA',21,'WR',25-28,'WYSSSK',39-60,
'AA',64-274 ##label USU
##cross-references GB:M34336
#note
the authors translated the codon AAG for residue 34 as
Arg
#note
this sequence has been revised in reference A45063
REFERENCE
A24011
#authors
Borchart, U.; Machleidt, W.; Schagger, H.; Link, T.A.; von
Jagow, G.
#journal
FEBS Lett. (1985) 191:125-130
#title
Isolation and amino acid sequence of the 8 kDa DCCD-binding
protein of beef heart ubiquinol:cytochrome c reductase.
#cross-references MUID:86030649
#accession A24011
##molecule_type protein
##residues 1-78 ##label BOR
#experimental_source heart
#note
53-Glu reacts rapidly and specifically with dicyclohexyl
carbodiimide
REFERENCE
S00003
#authors
Schagger, H.; Borchart, U.; Machleidt, W.; Link, T.A.; von
Jagow, G.
#journal
FEBS Lett. (1987) 219:161-168
#title
Isolation and amino acid sequence of the 'Rieske' iron sulfur
protein of beef heart ubiquinol:cytochrome c reductase.
#cross-references MUID:87247298
#accession S00003
##molecule_type protein
##residues 79-149,'A',151-268,'G',270-274 ##label SCH
REFERENCE
S14093
#authors
Cocco, T.; Lorusso, M.; Sardanelli, A.M.; Minuto, M.; Ronchi,
S.; Tedeschi, G.; Papa, S.
#journal
Eur. J. Biochem. (1991) 195:731-734
#title
Structural and functional characteristics of polypeptide
subunits of the bovine heart ubiquinol - cytochrome-c
reductase complex.
#cross-references MUID:91153313
#accession S14162
##molecule_type protein
##residues 79-90;142-149,'A',151-162 ##label COC
COMMENT
The transit peptide fragment remains bound in the cytochrome bcl
complex.
CLASSIFICATION
#superfamily ubiquinol--cytochrome-c reductase iron-sulfur
protein
KEYWORDS
2Fe-2S; acetylated amino end; electron transfer; iron-sulfur
protein; mitochondrion; oxidative phosphorylation;
oxidoreductase; respiratory chain; transmembrane protein
FEATURE
1-78
#domain transit peptide (mitochondrion) #status
experimental #label TNP\
1-78
#product ubiquinol--cytochrome-c reductase 8K protein
#status experimental #label M8K\
79-274
#product ubiquinol--cytochrome-c reductase iron-sulfur
protein #status experimental #label MAT\
112-129
#domain transmembrane #status predicted #label TM1\
132-150
#domain transmembrane #status predicted #label TM2\

1
#modified_site acetylated amino end (Met) (in
ubiquinol--cytochrome-c reductase 8K protein) #status
experimental\
217,222,236,238 #binding_site 2Fe-2S cluster (Cys) (covalent) #status
predicted
SUMMARY
#length 274 #molecular-weight 29547 #checksum 8317

Query Match 79.6%; Score 43; DB 1; Length 274;
Best Local Similarity 75.0%; Pred. No. 2.72e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 189 avevsqrlr 196
:| | | | |
QY 1 SVSVSQLR 8

RESULT 7
ENTRY S42366 #type complete
TITLE Lon proteinase homolog - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 07-Sep-1994 #sequence_revision 26-May-1995 #text_change
01-Sep-1995
ACCESSIONS S42366; S38500; S38860; S42365
REFERENCE S42365
#authors
Amerik, A.Y.; Petukhova, G.V.; Grigorenko, V.G.; Lykov, I.P.;
Varovoi, S.V.; Lipkin, V.M.; Gorbalenya, A.E.
#journal
FEBS Lett. (1994) 340:25-28
#title
Cloning and sequence analysis of cDNA for a human homolog of
eubacterial ATP-dependent Lon proteases.
#accession S42366
##molecule_type mRNA
##residues 1-937 ##label AME
##cross-references EMBL:X76040
##note
the authors translated the codon ATT for residue 39 as
Leu
SUMMARY
#length 937 #molecular-weight 103999 #checksum 915

Query Match 79.6%; Score 43; DB 13; Length 937;
Best Local Similarity 62.5%; Pred. No. 2.72e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 483 iavsqrlrg 490
:| | | | |
QY 2 VSVSQLRA 9

RESULT 8
ENTRY S57342 #type complete
TITLE ATP-dependent protease LON - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
03-Nov-1995
ACCESSIONS S57342
REFERENCE S57342
#authors
Wang, N.; Gottesman, S.; Willingham, M.C.; Gottesman, M.M.;
Maurizi, M.R.
#journal
Proc. Natl. Acad. Sci. U.S.A. (1993) 90:11247-11251
#title
A human mitochondrial ATP-dependent protease that is highly
homologous to bacterial Lon protease.
#accession S57342
##status preliminary
##molecule_type mRNA
##residues 1-962 ##label WAN
##cross-references EMBL:U02389
##note
nucleotide sequence is not given
SUMMARY
#length 962 #molecular-weight 106466 #checksum 3799

Query Match 79.6%; Score 43; DB 13; Length 962;
Best Local Similarity 62.5%; Pred. No. 2.72e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 508 iavsqrlrg 515
:| | | | |

```

```

QY      2 VSVSQLRA 9

RESULT  9
ENTRY   #type complete
TITLE   gamma-aminobutyrate permease - Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE    13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change

ACCESSIONS
REFERENCE S47017
#authors Metzger, E.; Halpern, Y.S.
#submission submitted to the EMBL Data Library, March 1992
#description Nucleotide sequence of the gapB gene of Escherichia coli K12
              encoding an integral membrane protein involved in
              gamma-aminobutyrate transport.
#accession S47017
#status preliminary
#molecule_type DNA
#residues 1-466 #label MET
#cross-references EMBL:X65104
SUMMARY #length 466 #molecular-weight 51017 #checksum 483

Query Match 77.8%; Score 42; DB 9; Length 466;
Best Local Similarity 71.4%; Pred. No. 4.31e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 378 lavsqr 384
   :|||||
QY 2 VSVSQLR 8

RESULT 10
ENTRY   #type complete
TITLE   pyruvate oxidoreductase - Entamoeba histolytica
ORGANISM #formal_name Entamoeba histolytica
DATE    13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change

ACCESSIONS
REFERENCE S58236
#authors Rodriguez, M.; Delgadillo, D.; Baez-Camargo, M.; Orozco, E.
#submission submitted to the EMBL Data Library, August 1995
#description Isolation of a pyruvate oxidoreductase gene from Entamoeba
              histolytica.
#accession S58236
#status preliminary
#molecule_type DNA
#residues 1-1165 #label ROD
#cross-references EMBL:X50193
SUMMARY #length 1165 #molecular-weight 128280 #checksum 1302

Query Match 77.8%; Score 42; DB 12; Length 1165;
Best Local Similarity 62.5%; Pred. No. 4.31e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 460 gvtvshlr 467
   :|:|||||
QY 1 SVSVSQLR 8

RESULT 11
ENTRY   #type complete
TITLE   hypothetical protein YMR229c - yeast (Saccharomyces
              cerevisiae)
ALTERNATE_NAMES hypothetical protein YM9959.11c
ORGANISM #formal_name Saccharomyces cerevisiae
DATE    19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change

ACCESSIONS
REFERENCE S57596
#authors Skelton, J.; Churcher, C.M.
#submission submitted to the EMBL Data Library, June 1995
#accession S57596

#molecule_type DNA
#residues 1-1729 #label SKE
#cross-references EMBL:X49939
#experimental_source strain AB972
GENETICS
#map_position 13R
SUMMARY #length 1729 #molecular-weight 193132 #checksum 3679

Query Match 77.8%; Score 42; DB 11; Length 1729;
Best Local Similarity 62.5%; Pred. No. 4.31e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 365 vnvshlrr 372
   :|:|||||
QY 2 VSVSQLRA 9

RESULT 12
ENTRY   #type complete
TITLE   gene 22 protein - human herpesvirus 3
ORGANISM #formal_name human herpesvirus 3, varicella-zoster virus
DATE    30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change

ACCESSIONS
REFERENCE D27343
#authors Davison, A.J.; Scott, J.E.
#journal J. Gen. Virol. (1986) 67:1759-1816
#title The complete DNA sequence of varicella-zoster virus.
#cross-references M01D:86306657
#accession D27343
#molecule_type DNA
#residues 1-2763 #label DAV
COMMENT The DNA sequence was obtained from EMBL, release 13.
GENETICS
#gene 22
CLASSIFICATION #superfamily varicella-zoster virus gene 22 protein
SUMMARY #length 2763 #molecular-weight 306339 #checksum 6847

Query Match 77.8%; Score 42; DB 4; Length 2763;
Best Local Similarity 66.7%; Pred. No. 4.31e+01;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1426 svtliselra 1434
   :|:|:|||||
QY 1 SVSVSQLRA 9

RESULT 13
ENTRY   #type complete
TITLE   plectin - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE    27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change

ACCESSIONS
REFERENCE A39638; S21876
#authors Wiche, G.; Becker, B.; Lubert, K.; Weitzer, G.; Castanon,
              M.J.; Hauptmann, R.; Stratowa, C.; Stewart, M.
#journal J. Cell Biol. (1991) 114:83-99
#title Cloning and sequencing of rat plectin indicates a 466-kD
              polypeptide chain with a three-domain structure based on a
              central alpha-helical coiled coil.
#cross-references M01D:91268156
#accession A39638
#status preliminary
#molecule_type mRNA
#residues 1-4140 #label WIC
#cross-references EMBL:X59601
SUMMARY #length 4140 #molecular-weight 464561 #checksum 2125

Query Match 77.8%; Score 42; DB 14; Length 4140;
Best Local Similarity 71.4%; Pred. No. 4.31e+01;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```



```
Db 1194 vtvtdlr 1200
QY 2 VSVSQLR 8

RESULT 14
ENTRY #type fragment
TITLE hypothetical protein 2 (3' of PTH270 promoter) - Streptomyces
        coelicolor (fragment)
ORGANISM #formal_name Streptomyces coelicolor
DATE 16-Feb-1994 #sequence_revision 01-Dec-1995 #text_change
ACCESSIONS D49692; S29938
REFERENCE A49692
#authors Tan, H.; Chater, K.F.
#journal J. Bacteriol. (1993) 175:933-940
#title Two developmentally controlled promoters of Streptomyces
        coelicolor A3(2) that resemble the major class of
        motility-related promoters in other bacteria.
#cross-references MUID:93163057
#contents A3(2)
#accession D49692
##status preliminary
##molecule_type DNA
##residues 1-111 ##label TAN
##cross-references NCBI:124681; NCBI:P124683
#note sequence extracted from NCBI backbone
REFERENCE S29937
#authors Bruton, C.J.
#submission submitted to the EMBL Data Library, October 1992
#accession S29938
##status preliminary
##molecule_type mRNA
##residues 1-111 ##label BRU
##cross-references EMBL:X58792
SUMMARY #length 111 #checksum 8772

Query Match 75.9%; Score 41; DB 10; Length 111;
Best Local Similarity 66.7%; Pred. No. 6.76e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 31 avsvkrlra 39
QY 1 SVSVSQLRA 9

RESULT 15
ENTRY #type complete
TITLE hypothetical protein YPL233w - yeast (Saccharomyces
        cerevisiae)
ALTERNATE_NAMES hypothetical protein P1401
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change
ACCESSIONS S61701; S65258; S65262
REFERENCE S61699
#authors Urrestarazu, L.A.
#submission submitted to the EMBL Data Library, December 1995
#accession S61701
##molecule_type DNA
##residues 1-216 ##label URR
##cross-references EMBL:X94561
REFERENCE S65251
#authors Urrestarazu, L.A.; Vissers, S.
#submission submitted to the Protein Sequence Database, May 1996
#accession S65258
##molecule_type DNA
##residues 1-216 ##label URW
##cross-references EMBL:Z73589
##experimental_source strain S288C (AB972)
REFERENCE S64899
#authors Pohl, T.M.
#submission submitted to the Protein Sequence Database, May 1996

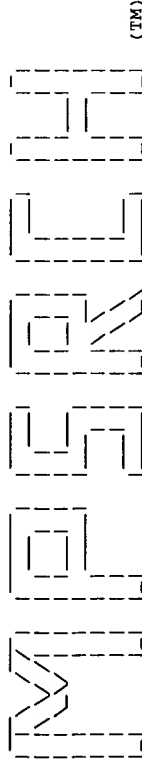
#accession S65262
##molecule_type DNA
##residues 1-84 ##label POH
##cross-references EMBL:Z73589
##experimental_source strain S288C (AB972)
GENETICS
#map_position 161
SUMMARY #length 216 #molecular-weight 25416 #checksum 4728

Query Match 75.9%; Score 41; DB 11; Length 216;
Best Local Similarity 62.5%; Pred. No. 6.76e+01;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 11 vtvqlrs 18
QY 2 VSVSQLRA 9

Search completed: Tue Jun 10 11:30:05 1997
Job time : 11 secs.
```

\*\*\*\*\*



\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:29:26 1997; MasPar time 2.03 Seconds  
Tabular output not generated. 93.823 Million cell updates/sec

Title: >US-08-231-565A-37  
Description: (1-9) from US08231565A.pep  
Perfect Score: 54  
Sequence: 1 SVSVSQLRA 9

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 22.211; Variance 20.944; scale 1.060

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	54	100.0	668	7	PM17_HUMAN PMEL 17 PROTEIN PRECU	9.03e-03
2	45	83.3	1199	6	NIFU_SYNY3 PYRUVATE-FLAVODOXIN O	2.34e+00
3	43	79.6	256	10	UCRI_RAT UBIQUINOL-CYTOCHROME	7.26e+00
4	43	79.6	274	10	UCRI_BOVIN UBIQUINOL-CYTOCHROME	7.26e+00
5	43	79.6	937	6	LONN_HUMAN MITOCHONDRIAL LON PRO	7.26e+00
6	43	79.6	982	6	LONN_HUMAN MITOCHONDRIAL LON PRO	7.26e+00
7	42	77.8	466	4	GABP_ECOLI GABA PERMEASE (4-AMIN	1.26e+01
8	42	77.8	2763	9	TEGU_VZVD LARGE TEGUMENT PROTEI	1.26e+01
9	42	77.8	4687	7	PLEC_RAT PLECTIN.	1.26e+01
10	41	75.9	328	9	SDHL_HUMAN L-SERINE DEHYDRATASE	2.16e+01
11	41	75.9	589	2	CAH_DUNSA CARBONIC ANHYDRASE (E	2.16e+01
12	41	75.9	753	11	YBU4_YEAST HYPOTHETICAL 86.4 KD	2.16e+01
13	41	75.9	958	10	VIA_CCMV 1A PROTEIN (CONTAINS:	2.16e+01
14	41	75.9	961	10	VIA_BMV 1A PROTEIN (CONTAINS:	2.16e+01
15	41	75.9	4451	4	GRSB_BACBR GRAMICIDIN S SYNTHETA	2.16e+01
16	40	74.1	93	5	IAPP_MOUSE ISLET AMYLOID POLYPEP	3.65e+01
17	40	74.1	93	5	IAPP_RAT ISLET AMYLOID POLYPEP	3.65e+01
18	40	74.1	148	8	R19E_METJA 30S RIBOSOMAL PROTEIN	3.65e+01
19	40	74.1	165	5	INI1_HUMAN INTERFERON-INDUCED 17	3.65e+01
20	40	74.1	167	10	VB03_VACCV PROTEIN B3.	3.65e+01
21	40	74.1	315	11	YX25_MYCTU PROBABLE INTEGRASE/RE	3.65e+01
22	40	74.1	671	1	ALYS_ENTFA AUTOLYSIN (EC 3.5.1.2	3.65e+01

RESULT	1	40	74.1	745	11	XY00_YEAST	HYPOTHETICAL 84.3 KD	3.65e+01
ID	PM17_HUMAN	STANDARD;				PURL_MYCTU	PHOSPHORIBOSYLFORMYLIG	3.65e+01
AC	P40967;					VG43_HSV11	HYPOTHETICAL GENE 43	3.65e+01
DT	01-FEB-1995 (REL. 31, CREATED)					NIFJ_KLEPN	PYRUVATE-FLAVODOXIN O	3.65e+01
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)					NIFJ_ANASP	PYRUVATE-FLAVODOXIN O	3.65e+01
DE	PMEL 17 PROTEIN PRECURSOR.					DYNA_RAT	DYNACTIN, 150 KD ISO	3.65e+01
GN	PMEL17.					SUF2_BACSU	SURFACTIN SYNTHETASE	3.65e+01
OS	HOMO SAPIENS (HUMAN).					3-DEHYDROQUINATE DEHY		3.65e+01
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;					AROQ_HSV11	HYPOTHETICAL 35.3 KD	6.11e+01
OC	EUTHERIA; PRIMATES.					YEHA_ECOLI	HYPOTHETICAL 36.9 KD	6.11e+01
RN	[1]					VVHA_VIBU	CYTOLYSIN PRECURSOR.	6.11e+01
RP	SEQUENCE FROM N.A.					REC_N_ECOLI	DNA REPAIR PROTEIN RE	6.11e+01
RX	MEDLINE; 92021023.					VP40_HSVBC	CAPSID PROTEIN P40 (C	6.11e+01
RA	KWON B.S., CHINTAMANENI C., KOZAK C.A., COPELAND N.G.,					VP40_HSVB	CAPSID PROTEIN P40 (V	6.11e+01
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,					BCL6_HUMAN	B-CELL LYMPHOMA 6 PRO	6.11e+01
RA	KIM K.-K.;					BCL6_MOUSE	B-CELL LYMPHOMA 6 PRO	6.11e+01
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).					ICD_CORGL	ISOCITRATE DEHYDROGEN	6.11e+01
CC	-I- FUNCTION: COULD BE A MELANOCYTIC ENZYME.					IDH_AZOVI	ISOCITRATE DEHYDROGEN	6.11e+01
CC	-I- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.					MCM3_XENLA	DNA REPLICATION LICEN	6.11e+01
CC	-I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).					NIRA_XENLA	NITROGEN ASSIMILATION	6.11e+01
CC	DR EMBL; M77348; G190106; -.					COPA_HUMAN	COATOMER ALPHA SUBUNI	6.11e+01
DR	MIM: 155550; -.					DNA2_YEAST	DNA REPLICATION HELIC	6.11e+01
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.					ZEPL_HUMAN	ZINC FINGER PROTEIN 4	6.11e+01
FT	SIGNAL	1	23					
FT	CHAIN	24	668				PMEL 17 PROTEIN.	
FT	TRANSMEM	575	595				POTENTIAL.	
FT	TRANSMEM	603	623				POTENTIAL.	
FT	DOMAIN	217	307				PKD.	
FT	DOMAIN	315	444				10 X 13 AA TANDEM REPEATS.	
FT	REPEAT	315	327				1.	
FT	REPEAT	328	340				2.	
FT	REPEAT	341	353				3.	
FT	REPEAT	354	366				4.	
FT	REPEAT	367	379				5.	
FT	REPEAT	380	392				6.	
FT	REPEAT	393	405				7.	
FT	REPEAT	406	418				8.	
FT	REPEAT	419	431				9.	
FT	REPEAT	432	444				10.	
FT	CARBOHYD	81	81				POTENTIAL.	
FT	CARBOHYD	106	106				POTENTIAL.	
FT	CARBOHYD	111	111				POTENTIAL.	
FT	CARBOHYD	321	321				POTENTIAL.	
FT	CARBOHYD	568	568				POTENTIAL.	
FT	SEQUENCE	668 AA;	70992 MW;				6E8E1AFO CRC32;	

ALIGNMENTS

```

Query Match      100.0%; Score 54; DB 7; Length 668;
Best Local Similarity 100.0%; Pred. No. 9.03e+03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 216 svsvsqira 224
    |||||
Qy 1 SVSVSQLRA 9

RESULT 2
ID NFEJ_SNNY3 STANDARD; PRT; 1199 AA.
AC P52965;
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE PYRUVATE-FLAVODOXIN OXIDOREDUCTASE (EC 1.-.-.-).
GN NFEJ OR SLL0741.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96127529.
RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
RA SUGIURA M., TABATA S.;
RL DNA RES. 2:153-166(1995).
CC -!- FUNCTION: OXIDOREDUCTASE REQUIRED FOR THE TRANSFER OF ELECTRONS
CC FROM PYRUVATE TO FLAVODOXIN, WHICH REDUCES NITROGENASE.
DR EMBL; D64005; G1006618; -
KW OXIDOREDUCTASE; NITROGEN FIXATION; ELECTRON TRANSPORT.
SQ SEQUENCE 1199 AA; 131457 MW; 7B134F88 CRC32;

Query Match      83.3%; Score 45; DB 6; Length 1199;
Best Local Similarity 75.0%; Pred. No. 2.34e+00;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 462 svtyshlr 469
    |||||
Qy 1 SVSVSQLR 8

RESULT 3
ID UCRI_RAT STANDARD; PRT; 256 AA.
AC P20788;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT PRECURSOR
DE (EC 1.10.2.2) (RIESKE IRON-SULFUR PROTEIN) (RISP) (FRAGMENT).
GN UQCRFSL.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89165853.
RA NISHIKIMI M., HOSOKAWA Y., TODA H., SUZUKI H., OZAWA T.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 159:19-25(1989).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -!- THE RIESKE PROTEIN IS A HIGH POTENTIAL 2FE-2S PROTEIN.
CC -!- CATALYTIC ACTIVITY: OH(2) + 2 FERRICYTOCHROME C = Q +
CC 2 FERROCYTOCHROME C.
CC -!- SUBUNIT: B-C1 COMPLEX CONTAINS 10 SUBUNITS; 3 RESPIRATORY
CC SUBUNITS, 2 CORE PROTEINS AND 5 LOW-MOLECULAR WEIGHT PROTEINS.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: TO RIESKE PROTEINS FROM OTHER SOURCES (MITOCHONDRIA,
CC BACTERIAL, CHLOROPLAST).
DR EMBL; M24542; G206681; -
DR PIR; A32296; A32296.

```

```

DR PROSITE; PS00199; RIESKE_1.
DR PROSITE; PS00200; RIESKE_2.
KW MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; IRON-SULFUR;
KW OXIDOREDUCTASE; INNER MEMBRANE; TRANSMEMBRANE; TRANSIT PEPTIDE.
FT NON_TER 1
FT TRANSIT <1 60 MITOCHONDRION.
FT CHAIN 61 256 RIESKE IRON-SULFUR PROTEIN.
FT METAL 199 199 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 201 201 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 218 218 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 221 221 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT DISULFID 204 220 BY SIMILARITY.
SQ SEQUENCE 256 AA; 27688 MW; E287CE7A CRC32;

Query Match      79.6%; Score 43; DB 10; Length 256;
Best Local Similarity 75.0%; Pred. No. 7.26e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 171 avevsqir 178
    :| |||||
Qy 1 SVSVSQLR 8

RESULT 4
ID UCRI_BOVIN STANDARD; PRT; 274 AA.
AC P13272; P07588;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE UBIQUINOL-CYTOCHROME C REDUCTASE IRON-SULFUR SUBUNIT PRECURSOR
DE (EC 1.10.2.2) (RIESKE IRON-SULFUR PROTEIN) (RISP) (CONTAINS:
DE UBIQUINOL-CYTOCHROME C REDUCTASE 8 KD PROTEIN (COMPLEX III SUBUNIT
DE IX)).
GN UQCRFSL.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HEART;
RX MEDLINE; 90211231.
RA USUI S., YU L., YU C.-A.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 167:575-579(1990).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RC TISSUE-HEART;
RX MEDLINE; 93231976.
RA BRANDT U., YU L., YU C.-A., TRUMPOWER B.L.;
RL J. BIOL. CHEM. 268:8387-8390(1993).
RN [3]
RP SEQUENCE OF 79-274.
RX MEDLINE; 87247298.
RA SCHAEGER H., BORCHART U., MACHLEIDT W., LINK T.A., VON JAGOW G.;
RL FEBS LETT. 219:161-168(1987).
RN [4]
RP SEQUENCE OF 1-78.
RC TISSUE-HEART;
RX MEDLINE; 86030649.
RA BORCHART U., MACHLEIDT W., SCHAGGER H., LINK T.A., VON JAGOW G.;
RL FEBS LETT. 191:125-130(1985).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 148-274.
RX MEDLINE; 96347356.
RA IWATA S., SAYNOVITS M., LINK T.A., MICHEL H.;
RL STRUCTURE 4:567-579(1996).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -!- FUNCTION: THE TRANSIT PEPTIDE OF THE RIESKE PROTEIN SEEMS TO FORM
CC PART OF THE B-C1 COMPLEX IN BEEF AND IS CONSIDERED TO BE THE
CC SUBUNIT IX OF THAT COMPLEX.
CC -!- THE RIESKE PROTEIN IS A HIGH POTENTIAL 2FE-2S PROTEIN.

```

```

CC CC -1- CATALYTIC ACTIVITY: QH(2) + 2 FERRICYTOCHROME C - Q +
CC CC 2 FERROCYTOCHROME C.
CC CC -1- SUBUNIT: BCL COMPLEX CONTAINS 10 SUBUNITS; 3 RESPIRATORY
CC CC SUBUNITS, 2 CORE PROTEINS AND 5 LOW-MOLECULAR WEIGHT PROTEINS.
CC CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC CC -1- SIMILARITY: TO RIESKE PROTEINS FROM OTHER SOURCES (MITOCHONDRIA,
CC CC BACTERIAL, CHLOROPLAST).
CC CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN BY EXTENSIVE
CC CC FRAMESHIFTS AND PROBABLE SEQUENCING ERRORS IN THE REGION OF THE
CC CC TRANSIT PEPTIDE (1-78).
CC DR EMBL; S58789; G299558; -.
CC DR EMBL; M34336; G163044; ALT_FRAME.
CC DR PIR; S00003; S00003.
CC DR PIR; A34660; A34660.
CC DR PIR; A24011; A24011.
CC DR PDB; IRIE; PRELIMINARY.
CC DR PROSITE; PS00199; RIESKE_1.
CC DR PROSITE; PS00200; RIESKE_2.
CC KW MITOCHONDRION; ELECTRON TRANSPORT; RESPIRATORY CHAIN; IRON-SULFUR;
CC KW OXIDOREDUCTASE; INNER MEMBRANE; TRANSMEMBRANE; TRANSIT PEPTIDE;
CC KW 3D-STRUCTURE.
CC FT CHAIN 1 78 UBIQUINOL-CYTOCHROME C REDUCTASE 8 KD
CC FT TRANSIT 1 78 MITOCHONDRION.
CC FT CHAIN 1 78 RIESKE IRON-SULFUR PROTEIN.
CC FT METAL 217 274 IRON-SULFUR (2FE-2S).
CC FT METAL 219 219 IRON-SULFUR (2FE-2S).
CC FT METAL 236 236 IRON-SULFUR (2FE-2S).
CC FT METAL 239 239 IRON-SULFUR (2FE-2S).
CC FT DISULFID 222 238 S -> A (IN REF. 3).
CC FT CONFLICT 150 150 D -> G (IN REF. 3).
CC FT CONFLICT 269 269
CC SQ SEQUENCE 274 AA; 29547 MW; EFDEACBC CRC32;

Query Match 79.6%; Score 43; DB 10; Length 274;
Best Local Similarity 75.0%; Pred. No. 7.26e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 189 avevsqrlr 196
:|||||
QY 1 SVSVSQLR 8

RESULT 5
ID LONN_HUMAN STANDARD; PRT; 937 AA.
AC P36777;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE MITOCHONDRIAL LON PROTEASE HOMOLOG PRECURSOR (EC 3.4.21.-).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
SEQUENCE FROM N.A.
TX TISSUE-BRAIN;
RX MEDLINE; 94068581.
RA WANG N.; GOTTESMAN S.; WILLINGHAM M.C.; GOTTESMAN M.M.; MAURIZI M.R.;
RL PROC. NATL. ACAD. SCI. U.S.A. 90:11247-11251(1993).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES. HIGHER LEVELS
CC FOUND IN HEART, BRAIN, LIVER AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S16; ALSO KNOWN AS THE
CC LON FAMILY OF ATP-DEPENDENT PROTEASES.
CC EMBL; U02389; G440874; -.
CC DR PROSITE; PS01046; LON_SER.
CC KW HYDROLASE; SERINE PROTEASE; ATP-BINDING; MITOCHONDRION;
CC KW TRANSIT PEPTIDE.
CC FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
CC FT CHAIN ? 962 MITOCHONDRIAL LON PROTEASE HOMOLOG.
CC FT NP_BIND 526 533 ATP (POTENTIAL).
CC FT ACT_SITE 858 858 BY SIMILARITY.
CC SQ SEQUENCE 962 AA; 106467 MW; E7A87F1F CRC32;

Query Match 79.6%; Score 43; DB 6; Length 962;
Best Local Similarity 62.5%; Pred. No. 7.26e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 508 lavsqrlr 515
:|||||
QY 2 VSVSQLRA 9

RESULT 7
ID GABP_ECOLI STANDARD; PRT; 466 AA.
AC P25527;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE GABA PERMEASE (4-AMINO BUTYRATE TRANSPORT CARRIER) (GABA-AMINOBUTYRATE
DE PERMEASE).
CC GN GABP.
CC OS ESCHERICHIA COLI.
CC OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
CC OC ENTEROBACTERIACEAE.
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-K12 / JM103;
CC RX MEDLINE; 94127927.
CC RA NIEGEMANN E.; SCHULZ A.; BARTSCH K.;
CC RL ARCH. MICROBIOL. 160:454-460(1993).
CC RN [2]

```

```

FT CHAIN ? 937 MITOCHONDRIAL LON PROTEASE HOMOLOG.
FT NP_BIND 501 508 ATP (POTENTIAL).
FT ACT_SITE 833 833 BY SIMILARITY.
SQ SEQUENCE 937 AA; 104000 MW; 43B79374 CRC32;

Query Match 79.6%; Score 43; DB 6; Length 937;
Best Local Similarity 62.5%; Pred. No. 7.26e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 483 lavsqrlr 490
:|||||
QY 2 VSVSQLRA 9

RESULT 6
ID LONN_HUMAN STANDARD; PRT; 962 AA.
AC P36776;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE MITOCHONDRIAL LON PROTEASE HOMOLOG PRECURSOR (EC 3.4.21.-).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
CC RN [1]
SEQUENCE FROM N.A.
TX TISSUE-BRAIN;
RX MEDLINE; 94068581.
RA WANG N.; GOTTESMAN S.; WILLINGHAM M.C.; GOTTESMAN M.M.; MAURIZI M.R.;
RL PROC. NATL. ACAD. SCI. U.S.A. 90:11247-11251(1993).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES. HIGHER LEVELS
CC FOUND IN HEART, BRAIN, LIVER AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S16; ALSO KNOWN AS THE
CC LON FAMILY OF ATP-DEPENDENT PROTEASES.
CC EMBL; U02389; G440874; -.
CC DR PROSITE; PS01046; LON_SER.
CC KW HYDROLASE; SERINE PROTEASE; ATP-BINDING; MITOCHONDRION;
CC KW TRANSIT PEPTIDE.
CC FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
CC FT CHAIN ? 962 MITOCHONDRIAL LON PROTEASE HOMOLOG.
CC FT NP_BIND 526 533 ATP (POTENTIAL).
CC FT ACT_SITE 858 858 BY SIMILARITY.
CC SQ SEQUENCE 962 AA; 106467 MW; E7A87F1F CRC32;

Query Match 79.6%; Score 43; DB 6; Length 962;
Best Local Similarity 62.5%; Pred. No. 7.26e+00;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 508 lavsqrlr 515
:|||||
QY 2 VSVSQLRA 9

RESULT 7
ID GABP_ECOLI STANDARD; PRT; 466 AA.
AC P25527;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE GABA PERMEASE (4-AMINO BUTYRATE TRANSPORT CARRIER) (GABA-AMINOBUTYRATE
DE PERMEASE).
CC GN GABP.
CC OS ESCHERICHIA COLI.
CC OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
CC OC ENTEROBACTERIACEAE.
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-K12 / JM103;
CC RX MEDLINE; 94127927.
CC RA NIEGEMANN E.; SCHULZ A.; BARTSCH K.;
CC RL ARCH. MICROBIOL. 160:454-460(1993).
CC RN [2]

```



Db 1741 vvtqlr 1747  
QY 2 VSVSQR 8  
RESULT 10  
ID SDHL HUMAN STANDARD; PRT; 328 AA.  
AC P20132.  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)  
DE L-SERINE DEHYDRATASE (EC 4.2.1.13) (L-SERINE DEAMINASE).  
GN SDH.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE; 89380167.  
RA OGAWA H., GOMI T., KONISHI K., DATE T., NAKASHIMA H., NOSE K.,  
RA MATSUDA Y., PERAINO C., PITOT H.C., FUJIOKA M.;  
RL J. BIOL. CHEM. 264:15818-15823(1989).  
CC -!- CATALYTIC ACTIVITY: L-SERINE + H(2)O = PYRUVATE + NH(3) + H(2)O.  
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -!- PATHWAY: GLUCONEOGENESIS FROM SERINE.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: TO THREONINE DEHYDRATASES BIOSYNTHETIC AND CATABOLIC.  
DR EMBL; J05037; G338030; -.  
DR PIR; A34232; DWHUT.  
DR MIM; 182128; -.  
DR PROSITE; P500165; DEHYDRATASE\_SER\_THR.  
KW LYSASE; PYRIDOXAL PHOSPHATE; GLUCONEOGENESIS; LIVER.  
FT BINDING 41 41 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
FT DOMAIN 163 176 REGION OF ACTIVE SITE (POTENTIAL).  
SQ SEQUENCE 328 AA; 34702 MW; 0B6F1251 CRC32;  
Query Match 75.9%; Score 41; DB 9; Length 328;  
Best Local Similarity 55.6%; Pred. No. 2.16e+01;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
Db 307 nlsaqrla 315  
QY 1 SVSVSOLRA 9  
RESULT 11  
ID CAH\_DUNSA STANDARD; PRT; 589 AA.  
AC P54212;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE CARBONIC ANHYDRASE (EC 4.2.1.1) (CARBONATE DEHYDRATASE).  
GN DCA.  
OS DUNALIELLA SALINA.  
OC EUKARYOTA; PLANTA; PHYCOPHYTA; CHLOROPHYTA (GREEN ALGAE);  
OC CHLOROPHYCEAE; VOLVOCALES; DUNALIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96279304.  
RA FISHER M., GORHAM I., PICK U., ZAMIR A.;  
RL J. BIOL. CHEM. 271:17718-17723(1996).  
CC -!- FUNCTION: REVERSIBLE HYDRATATION OF CARBON DIOXIDE.  
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.  
CC -!- INDUCTION: BY SALT.  
CC -!- SIMILARITY: TO OTHER CARBONIC ANHYDRASES.  
DR EMBL; U53811; G1431878; -.  
KW LYASE; ZINC.  
FT DOMAIN 390 589 CATALYTIC.  
FT METAL 418 418 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 420 420 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 440 440 ZINC (CATALYTIC) (BY SIMILARITY).

SQ SEQUENCE 589 AA; 64257 MW; 2D375843 CRC32;  
Query Match 75.9%; Score 41; DB 2; Length 589;  
Best Local Similarity 62.5%; Pred. No. 2.16e+01;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Db 276 tvsvaqlk 283  
QY 1 SVSVSQR 8  
RESULT 12  
ID YBU4\_YEAST STANDARD; PRT; 753 AA.  
AC P38254;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 86.4 KD PROTEIN IN PHO5-VPS15 INTERGENIC REGION.  
GN YBPO94W OR YBRO821.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE; 95208357.  
RA MANNHAUPT G., STUCKA R., EHNLE S., VETTER I., FELDMANN H.;  
RL YEAST 10:1363-1381(1994).  
RN [2]  
RP SEQUENCE OF 167-753 FROM N.A.  
RC STRAIN=2180;  
RA DEKKER P.J.T., HOEKERT W., VAN OOSTERUM K., GRIVELL L.A.;  
RL SUBMITTED (DEC-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: TO S.POMBE SPAC12B10.04.  
CC -!- SIMILARITY: SOME, TO TUBULIN-TYROSINE LIGASE.  
DR EMBL; X78993; G476052; -.  
DR EMBL; Z35963; G536367; -.  
DR EMBL; X69881; G4088; -.  
DR PIR; S44676; S44676.  
KW HYPOTHETICAL PROTEIN.  
FT CONFLICT 352 355 HALK -> TPE (IN REF. 2).  
FT CONFLICT 450 450 R -> A (IN REF. 2).  
FT CONFLICT 563 563 H -> R (IN REF. 2).  
SQ SEQUENCE 753 AA; 86438 MW; CE6A6B5D CRC32;  
Query Match 75.9%; Score 41; DB 11; Length 753;  
Best Local Similarity 71.4%; Pred. No. 2.16e+01;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Db 534 vnlsqrl 540  
QY 2 VSVSQR 8  
RESULT 13  
ID VIA\_CCMV STANDARD; PRT; 958 AA.  
AC P27752;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE 1A PROTEIN (CONTAINS: HELICASE AND METHYLTRANSFERASE).  
OS COWPEA CHLOROTIC MOTTLE VIRUS (CCMV).  
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; BROMOVIRIDAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92074218.  
RA DEIANOTT A.M., BUJARSKI J.J.;  
RL VIROLOGY 185:553-562(1991).  
CC -!- FUNCTION: MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS.  
CC -!- CONTAINS AN HELICASE DOMAIN AND A METHYLTRANSFERASE DOMAIN. THE  
CC METHYLTRANSFERASE DOMAIN IS PROBABLY INVOLVED IN VIRAL RNA  
CC CAPPING.  
CC -!- SIMILARITY: TO 1A PROTEIN FROM BMV, CMV, PSV AND TAV.  
DR EMBL; M65139; G331599; -.

DR PIR: B41699; PIBVCC.  
 KW HELICASE; ATP-BINDING; TRANSFERASE; METHYLTRANSFERASE.  
 FT NP\_BIND 682 689 ATP (POTENTIAL).  
 SQ SEQUENCE 958 AA; 109138 MW; 10F9C575 CRC32;

Query Match 75.9%; Score 41; DB 10; Length 958;  
 Best Local Similarity 55.6%; Pred. No. 2.16e+01;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 477 svp1sr1rs 485

QY 1 SVSSQLRA 9

||||:|

RESULT 14  
 ID VIA\_BW STANDARD; PRT; 961 AA.  
 AC P03588;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)  
 DE 1A PROTEIN (CONTAINS: HELICASE AND METHYLTRANSFERASE).  
 OS BROME MOSAIC VIRUS (BMV).  
 OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; BROMOVIRIDAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 84114904.  
 RA AHLQUIST P., DASGUPTA R., KAESBERG P.;  
 RL J. MOL. BIOL. 172:369-383(1984).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE REPLICATION OF THE VIRUS.  
 CC CONTAINS AN HELICASE DOMAIN AND A METHYLTRANSFERASE DOMAIN. THE  
 CC METHYLTRANSFERASE DOMAIN IS PROBABLY INVOLVED IN VIRAL RNA  
 CC CAPPING.  
 CC -!- SIMILARITY: TO 1A PROTEIN FROM CCMV, CMV, PSV AND TAV.  
 DR EMBL; X03380; G58729; -.  
 DR PIR; A04196; PIBVA.  
 KW HELICASE; ATP-BINDING; TRANSFERASE; METHYLTRANSFERASE.  
 FT NP\_BIND 685 692 ATP (POTENTIAL).  
 SQ SEQUENCE 961 AA; 109209 MW; 11FC15E2 CRC32;

Query Match 75.9%; Score 41; DB 10; Length 961;  
 Best Local Similarity 55.6%; Pred. No. 2.16e+01;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 476 svp1sr1rt 484

QY 1 SVSSQLRA 9

||||:|

RESULT 15  
 ID GRSB\_BACBR STANDARD; PRT; 4451 AA.  
 AC P14688;  
 DT 01-APR-1990 (REL. 14, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE GRAMICIDIN S SYNTHETASE II (GRAMICIDIN S BIOSYNTHESIS GRSB PROTEIN)  
 DE (EC 6.-.-.-).  
 GN GRSB OR GRS2.  
 OS BACILLUS BREVIS.  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 9999;  
 RX MEDLINE; 92219998.  
 RA TURGAY K., KRAUSE M., MARAHIEL M.A.;  
 RL MOL. MICROBIOL. 6:529-546(1992).  
 RN [2]  
 RP SEQUENCE OF 1-948 FROM N.A.  
 RC STRAIN-NAGANO;  
 RX MEDLINE; 92041751.  
 RA HORI K., YAMAMOTO Y., TOKITA K., SAITO F., KUROTSU T., KANDA M.,  
 RA OKAMURA K., FURUYAMA J., SAITO Y.,  
 RL J. BIOCHEM. 110:111-119(1991).  
 RN [3]

RP SEQUENCE OF 1-143 FROM N.A.  
 RC STRAIN-ATCC 9999;  
 RX MEDLINE; 90008776.  
 RA KRAETZSCHMAR J., KRAUSE M., MARAHIEL M.A.;  
 RL J. BACTERIOL. 171:5422-5429(1989).  
 RN [4]  
 RP SEQUENCE OF 1-15, AND CHARACTERIZATION.  
 RC STRAIN-NAGANO;  
 RX MEDLINE; 92011463.  
 RA KUROTSU T., HORI K., KANDA M., SAITO Y.;  
 RL J. BIOCHEM. 109:763-769(1991).  
 CC -!- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME, ABLE TO  
 CC ACTIVATE AND POLYMERIZE THE AMINO ACIDS PRO, VAL, ORN AND LEU.  
 CC ACTION SITES FOR THESE AA CONSIST OF INDIVIDUAL DOMAINS.  
 CC -!- PATHWAY: NON-RIBOSOMAL BIOSYNTHESIS OF THE CYCLIC PEPTIDE  
 CC ANTIBIOTIC GRAMICIDIN S (D-PHE-PRO-VAL-ORN-LEU)2.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- COFACTOR: CONTAINS FOUR COVALENTLY BOUND PHOSPHOPANTHETHEINES.  
 CC -!- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
 DR EMBL; X61658; G39372; -.  
 DR EMBL; M29703; G143031; -.  
 DR EMBL; X15577; G39370; -.  
 DR EMBL; D00938; G216276; -.  
 DR PIR; S20542; YGBSG2.  
 DR HSSP; P02901; IACP.  
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE.  
 DR PROSITE; PS00455; AMP\_BINDING.  
 KW LIGASE; ANTIBIOTIC BIOSYNTHESIS; MULTIFUNCTIONAL ENZYME;  
 KW REPEAT; PHOSPHOPANTHETHEINE.  
 FT INIT\_MET 0  
 FT REPEAT 466 1043  
 FT REPEAT 1521 2080  
 FT REPEAT 2538 3135  
 FT REPEAT 3591 4173  
 FT BINDING 1005 1005  
 FT BINDING 2041 2041  
 FT BINDING 3087 3087  
 FT BINDING 4125 4125  
 FT CONFLICT 274 274  
 FT CONFLICT 418 418  
 FT CONFLICT 654 664  
 FT CONFLICT 941 946  
 FT CONFLICT 941 946  
 SQ SEQUENCE 4451 AA; 510036 MW; F85BCDFC CRC32;  
 DOMAIN 1 (PROLINE-ACTIVATING).  
 DOMAIN 2 (VALINE-ACTIVATING).  
 DOMAIN 3 (ORNITHINE-ACTIVATING).  
 DOMAIN 4 (LEUCINE-ACTIVATING).  
 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 PHOSPHOPANTHETHEINE (BY SIMILARITY).  
 H -> D (IN REF. 2).  
 K -> R (IN REF. 2).  
 NAVLTCTVTKKF -> TCSFDVCYQEI (IN REF. 2).  
 HVRHL -> QLPLTP (IN REF. 2).

Query Match 75.9%; Score 41; DB 4; Length 4451;  
 Best Local Similarity 62.5%; Pred. No. 2.16e+01;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 2997 vnvaelra 3004

QY 2 VSVSQLRA 9

||||:|

Search completed: Tue Jun 10 11:29:36 1997  
 Job time : 10 secs.

\*\*\*\*\*  
[W][A][T][E][R][M][A][N]  
[P][R][O][T][E][I][N]  
[S][E][Q][U][E][N][C][E]  
[I][N]  
[T][E][R][M][I][N][A][T][I][O][N]  
(TM)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jun 10 11:31:45 1997; MasPar time 2.01 Seconds  
Tabular output not generated. 48.867 Million cell updates/sec

Title: >US-08-231-565A-38  
Description: (1-9) from US08231565A.pep  
Perfect score: 65  
Sequence: 1 YLEPGPVTA 9

Scoring table:  
PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part3 3:part4 4:part5 5:part6 6:part7 7:part8  
8:part9 9:part10 10:part11 11:part12 12:part13 13:part14  
14:part15 15:part16 16:part17 17:part18 18:part19  
19:part19

Statistics: Mean 16.492; Variance 41.546; scale 0.397

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	65	100.0	9 15 R84206	gp100 melanoma antige	1.56e-01
2	65	100.0	9 15 R82098	Melanoma-specific imm	1.56e-01
3	65	100.0	661 14 R78646	Melanoma associated a	1.56e-01
4	65	100.0	661 15 R84855	MART-1 melanoma antig	1.56e-01
5	65	100.0	661 15 R84854	MART-1 melanoma antig	1.56e-01
6	61	93.8	9 15 R84839	Modified melanocyte-m	5.46e-01
7	61	93.8	9 15 R84838	Modified melanocyte-m	5.46e-01
8	61	93.8	9 15 R84837	Modified melanocyte-m	5.46e-01
9	61	93.8	9 15 R82106	Melanoma-specific mut	5.46e-01
10	60	92.3	9 15 R82105	Melanoma-specific mut	7.45e-01
11	60	92.3	9 15 R84840	Modified melanocyte-m	7.45e-01
12	59	90.8	9 15 R82107	Melanoma-specific imm	1.01e+00
13	59	90.8	9 15 R84847	Modified melanocyte-m	1.01e+00
14	58	89.2	9 15 R84846	Modified melanocyte-m	1.38e+00
15	58	89.2	9 15 R82104	Melanoma-specific mut	1.38e+00
16	57	87.7	9 15 R82103	Melanoma-specific mut	1.87e+00
17	57	87.7	9 15 R82100	Melanoma-specific mut	1.87e+00
18	57	87.7	9 15 R82102	Melanoma-specific mut	1.87e+00
19	57	87.7	9 15 R82101	Melanoma-specific mut	1.87e+00
20	55	84.6	9 15 R84845	Modified melanocyte-m	3.43e+00

21	55	84.6	9 15 R84850	Modified melanocyte-m	3.43e+00
22	54	83.1	9 15 R84849	Modified melanocyte-m	4.63e+00
23	53	81.5	9 15 R84842	Modified melanocyte-m	6.24e+00
24	53	81.5	9 15 R82099	Melanoma-specific mut	6.24e+00
25	53	81.5	9 15 R84841	Modified melanocyte-m	6.24e+00
26	51	78.5	9 15 R84848	Modified melanocyte-m	1.13e+01
27	51	78.5	9 15 R84844	Modified melanocyte-m	1.13e+01
28	50	76.9	3910 12 R66462	ALL-1 (acute lymphocy	1.51e+01
29	50	76.9	3910 7 R38470	ALL-1 protein.	1.51e+01
30	50	76.9	3959 10 R52971	Product of the cDNA e	1.51e+01
31	49	75.4	9 15 R84851	Modified melanocyte-m	2.03e+01
32	48	73.8	9 15 R84843	Modified melanocyte-m	2.71e+01
33	47	72.3	9 15 R84852	Modified melanocyte-m	3.61e+01
34	47	72.3	281 1 P90498	STX pillin of Bordetel	3.61e+01
35	47	72.3	1713 13 R70148	Deduced sequence of c	3.61e+01
36	45	69.2	534 15 R85386	Barbary duck parvovir	6.38e+01
37	45	69.2	732 15 R85385	Barbary duck parvovir	6.38e+01
38	45	69.2	3457 12 R62504	Large polyprotein seq	6.38e+01
39	44	67.7	9 15 R84853	Modified melanocyte-m	8.44e+01
40	44	67.7	109 2 R11904	Human nm23 protein.	8.44e+01
41	43	66.2	152 2 R11903	Human nm23 protein.	1.12e+02
42	43	66.2	482 13 R75386	Natural resistance-as	1.12e+02
43	43	66.2	584 13 R75387	Natural resistance-as	1.12e+02
44	43	66.2	3898 2 R06996	Protein characteristi	1.12e+02
45	43	66.2	3898 11 R60543	55 kilodalton protein	1.12e+02

ALIGNMENTS

RESULT 1  
ID R84206 standard; Peptide; 9 AA.  
AC R84206;  
DT 25-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G9-280).  
KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI; 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 55; Page 131; 184pp; English.  
CC The immunogenic peptide is derived from cDNA25 (R84854), a  
CC melanoma antigen derivative of gp100 (see R84855). The  
CC peptide and its derivatives (see R84200-R84211) are used in  
CC medicaments (vaccines) for the treatment or prevention (by  
CC immunization) of melanoma. Antibodies against melanoma-specific  
CC antigens and its immunogenic peptides may be used in the  
CC detection and isolation of the antigen from a sample, the  
CC detection of which is indicative of a disease state  
CC (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 100.0%; Score 65; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.56e-01;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 ylepgpvta 9  
QY 1 YLEPGPVTA 9  
RESULT 2  
ID R82098 standard; peptide; 9 AA.





```

FT /label= antigenic.peptide
FT /note= "see R84199"
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
DR N-PSDB; T02716.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Claim 81; Fig 5A; 184pp; English.
CC cDNA2 is a melanoma antigen (MART-1) which is recognized by
CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-
CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source
CC of immunogenic peptides (see R84199) which are optionally modified
CC (see R84200-R84211) to enhance their binding to a MHC molecule and
CC used in medicaments, especially vaccines, for the treatment or
CC prevention (by immunisation) of melanoma. Antibodies against cDNA2
CC and its immunogenic peptides may be used in the detection and
CC isolation of the antigen from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 661 AA;

Query Match 100.0%; Score 65; DB 15; Length 661;
Best Local Similarity 100.0%; Pred. No. 1.56e-01;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 280 ylepgpvt 288
| | | | | | | |
QY 1 YLEPGPVT 9

RESULT 6
ID R84839;
AC R84839;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-9L.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 93.8%; Score 61; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.46e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ylepgpvt 8
| | | | | | | |
QY 1 YLEPGPVT 8

RESULT 8
ID R84837;
AC R84837;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-9V.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 93.8%; Score 61; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.46e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ylepgpvt 8
| | | | | | | |
QY 1 YLEPGPVT 8

RESULT 8
ID R84837;
AC R84837;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-9L.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

```

```

QY 1 YLEPGPVT 8

RESULT 7
ID R84838;
AC R84838;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-9L.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 93.8%; Score 61; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.46e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ylepgpvt 8
| | | | | | | |
QY 1 YLEPGPVT 8

RESULT 8
ID R84837;
AC R84837;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-9V.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN W09529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

```

```

Query Match      93.8%; Score 61; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 5,46e-01;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db      1 ylepgpvt 8
        |
QY      1 YLEPGPVT 8
        |

RESULT 9
ID R82106 standard; peptide; 9 AA.
AC R82106;
DT 25-MAR-1996 (first entry)
DE Melanoma-specific mutant immunogen epitope peptide.
KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;
KW cytotoxic T cell; lymphocyte; HLA-A2.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc_difference 9
FT /label= Thr, Ser, Pro, Gly
PN WO9522561-A2.
PD 24-AUG-1995.
PF 16-FEB-1995; U01991.
PR 16-FEB-1994; US-197399.
PR 29-APR-1994; US-234784.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;
PI WPI; 95-302688/39.
PT Melanoma-specific immunogen comprises epitope(s) homologous with
PT pMel.17 - are highly potent stimulators of HLA-A2+CTL's useful in
PT adoptive immuno-therapy
PS Claim 8; Page 130; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 comprises one
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-
CC R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines.
SQ Sequence 9 AA;

Query Match      92.3%; Score 60; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 7,45e-01;
Matches      8; Conservative      0; Mismatches      1; Indels      0; Gaps      0;

Db      1 ylepgpvt 9
        |
QY      1 YLEPGPVT 9
        |

RESULT 11
ID R84840 standard; Peptide; 9 AA.
AC R84840;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-1F.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
PI WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match      92.3%; Score 60; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 7,45e-01;
Matches      8; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

Db      1 ylepgpvt 9
        |
QY      1 YLEPGPVT 9
        |

RESULT 12
ID R82107 standard; peptide; 9 AA.
AC R82107;
DT 25-MAR-1996 (first entry)
DE Melanoma-specific immunogen epitope peptide, 946I.
KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;
KW cytotoxic T cell; lymphocyte; HLA-A2.
OS Homo sapiens.
PN WO9522561-A2.
PD 24-AUG-1995.
PF 16-FEB-1995; U01991.
PR 16-FEB-1994; US-197399.

```

```

PR 29-APR-1994; US-234784.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;
DR WPI; 95-302688/39.
PT Melanoma-specific immunogen comprises epitope(s) homologous with
PT pMel-17 - are highly potent stimulators of HLA-A2+CTL's useful in
PT adoptive immuno-therapy
PS Claim 10; Page 90; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 comprises one
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-
CC R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines.
SQ Sequence 9 AA;

Query Match 90.8%; Score 59; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.01e+00;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 ylepqpvt 9
   |||||||
QY 1 YLEPGPVT 9

RESULT 13
ID R84847 standard; Peptide; 9 AA.
AC R84847;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-3A.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;
KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 90.8%; Score 59; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.01e+00;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 ylapgpvt 9
   |||||||
QY 1 YLEPGPVT 9

RESULT 14
ID R84846 standard; Peptide; 9 AA.
AC R84846;
DT 25-APR-1996 (first entry)
DE Modified melanocyte-melanoma specific antigenic peptide G9-280-3S.
KW MART-1; M9-2; melanoma antigen recognised by T-cells; melanoma;

```

```

KW metastatic melanoma; tumour-associated antigen;
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;
KW therapy; vaccine.
OS Synthetic.
PN WO9529193-A2.
PD 02-NOV-1995.
PF 21-APR-1995; U05063.
PR 22-APR-1994; US-231565.
PR 05-APR-1995; US-417174.
PA (USSH ) US SEC DEPT HEALTH.
PI Kawakami Y, Rosenberg SA;
DR WPI; 95-382963/49.
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also
PT vectors, host cells and antibodies, used to detect, treat and
PT immunise animal against melanoma.
PS Example 5; Page 108; 184pp; English.
CC R84837-853 are G9-280 peptides modified to improve immunogenicity.
CC G9-280 is an immunogenic peptide based on the melanoma derived antigen,
CC gp100 (see R84206). The peptides are used in medicaments for the
CC treatment or prevention (by immunization) of melanoma. Antibodies
CC against MART-1 and its immunogenic peptides may be used in the detection
CC and isolation of MART-1 from a sample, the detection of which is
CC indicative of a disease state (melanoma or metastatic melanoma).
SQ Sequence 9 AA;

Query Match 89.2%; Score 58; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.38e+00;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 ylsppgvt 9
   |||||||
QY 1 YLEPGPVT 9

RESULT 15
ID R82104 standard; peptide; 9 AA.
AC R82104;
DT 25-MAR-1996 (first entry)
DE Melanoma-specific mutant immunogen epitope peptide.
KW Melanoma; immunogen; epitope; homologue; vaccine; immunotherapy;
KW cytotoxic T cell; lymphocyte; HLA-A2.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 7
FT /label= Ile, Leu, Met
PN WO9522561-A2.
PD 24-AUG-1995.
PF 16-FEB-1995; U01991.
PR 16-FEB-1994; US-197399.
PR 29-APR-1994; US-234784.
PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
PI Cox AL, Engelhard VH, Hunt DF, Shabanowitz J, Slingluff CL;
DR WPI; 95-302688/39.
PT Melanoma-specific immunogen comprises epitope(s) homologous with
PT pMel-17 - are highly potent stimulators of HLA-A2+CTL's useful in
PT adoptive immuno-therapy
PS Claim 8; Page 130; 148pp; English.
CC A melanoma-specific immunogen homologous with pMel-17 comprises one
CC or more CTL (cytotoxic T lymphocyte) epitopes from the group R82098-
CC R82194 capable of eliciting a CTL response. The epitopes R82098-
CC R82108 are of particular interest. The immunogen can be used for
CC partial protection in mammals against melanoma peptides which are
CC homologous with pMel-17 are highly potent stimulators of HLA-A2+
CC CTLs in several cell lines and can be used in immunotherapy or
CC incorporated into immunogenic conjugates as vaccines.
SQ Sequence 9 AA;

Query Match 89.2%; Score 58; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.38e+00;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 ylepqpvt 9
   |||||||
QY 1 YLEPGPVT 9

```

Tue Jun 10 15:07:05 1997

US-08-231-565A-38.rag

Page 6

QY 1 YLEPCPVT 9

Search completed: Tue Jun 10 11:31:53 1997  
Job time : 8 secs.

\*\*\*\*\*

W P E R E H (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:31:17 1997; MasPar time 2.67 Seconds  
95.995 Million cell updates/sec  
Tabular output not generated.

Title: >US-08-231-565A-38  
Description: (1-9) from US08231565A.pap  
Perfect Score: 65  
Sequence: 1 YLEPGPVT A 9

Scoring table: PAM 150  
Gap 15

Searched: 89912 seqs, 28507787 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir50

1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 22.260; Variance 27.785; scale 0.801

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	65	100.0	661	13	glycoprotein gp100 P	1.13e-03
2	65	100.0	668	13	melanocyte-specific	1.13e-03
3	58	89.2	491	14	melanoma antigen hom	4.80e-02
4	55	84.6	3866	14	All-1 protein -GFE f	2.23e-01
5	55	84.6	3869	14	All-1 protein +GFE f	2.23e-01
6	51	78.5	626	14	Pmel 17 protein - mo	1.61e+00
7	50	76.9	3910	13	trithorax homolog HT	2.61e+00
8	50	76.9	3968	13	trithorax homolog HT	2.61e+00
9	49	75.4	201	8	matrix protein M2 -	4.19e+00
10	47	72.3	201	7	fimbrial protein pre	1.06e+01
11	47	72.3	201	7	fimbrial protein pre	1.06e+01
12	47	72.3	207	7	fimbrial protein pre	1.06e+01
13	47	72.3	208	7	fimbrial protein pre	1.06e+01
14	47	72.3	1234	12	I factor 2 (transpos	1.06e+01
15	47	72.3	1713	13	adhesive ligand epil	1.06e+01
16	47	72.3	2257	5	DNA-directed DNA pol	1.06e+01
17	46	70.8	747	5	AMP deaminase (EC 3	1.67e+01
18	46	70.8	1010	13	F11 protein - chicke	1.67e+01
19	46	70.8	1091	13	contactin precursor	1.67e+01
20	46	70.8	1174	13	protein-tyrosine-pho	1.67e+01
21	46	70.8	1175	5	protein-tyrosine-pho	1.67e+01

22 46 70.8 1176 14 I58345 protein tyrosine pho 1.67e+01  
23 46 70.8 26926 13 I38344 titin - human 2.61e+01  
24 45 69.2 C64144 hypothetical protein 2.61e+01  
25 45 69.2 501 4 P1WLEP L1 protein - Europea 2.61e+01  
26 45 69.2 513 4 P1WLEP L1 protein - deer pa 2.61e+01  
27 45 69.2 732 16 S52210 VPL protein - muscov 2.61e+01  
28 44 67.7 152 14 A38369 nucleoside-diphospha 4.06e+01  
29 44 67.7 152 14 S29441 nucleoside-diphospha 4.06e+01  
30 44 67.7 152 13 I59581 c-myc transcription 4.06e+01  
31 44 67.7 157 14 S58038 probable olfactory r 4.06e+01  
32 44 67.7 157 13 S57998 probable olfactory r 4.06e+01  
33 44 67.7 157 13 S58014 probable olfactory r 4.06e+01  
34 44 67.7 176 13 A49798 tumor metastasis inh 4.06e+01  
35 44 67.7 504 4 P1W51 L1 protein - human p 4.06e+01  
36 44 67.7 602 11 S46676 hypothetical protein 4.06e+01  
37 43 66.2 152 14 I52807 protein nm23 - mouse 6.26e+01  
38 43 66.2 152 14 A45208 nucleoside diphospha 6.26e+01  
39 43 66.2 152 14 B33386 Nm23 protein - mouse 6.26e+01  
40 43 66.2 166 14 A46557 tumor metastasis inh 6.26e+01  
41 43 66.2 422 16 S27925 gene Lf1 protein - h 6.26e+01  
42 43 66.2 468 16 S41671 gene wg protein - fr 6.26e+01  
43 43 66.2 481 5 B43674 protein kinase (EC 2 6.26e+01  
44 43 66.2 1043 14 B33754 recombination-activa 6.26e+01  
45 43 66.2 3898 4 GNVVHC genome polyprotein - 6.26e+01

## ALIGNMENTS

RESULT 1  
ENTRY A53668 #type complete  
TITLE glycoprotein gp100 precursor, melanocyte lineage - human  
ALTERNATE\_NAMES melanoma antigen 25  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 25-May-1996  
ACCESSIONS A53668; A55753  
REFERENCE A53668  
#authors Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Fidor, C.G.  
#journal J. Biol. Chem. (1994) 269:20126-20133  
#title Molecular characterization of the melanocyte lineage-specific antigen gp100.  
#accession A53668  
#molecule\_type mRNA  
#residues 1-661 #label ADE  
REFERENCE A55753  
#authors Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Yanneli, J.R.; Adema, G.J.; Miki, T.; Rosenberg, S.A.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1994) 91:6458-6462  
#title Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes associated with in vivo tumor rejection.  
#accession A55753  
#status nucleic acid sequence not shown; not compared with conceptual translation  
#molecule\_type mRNA  
#residues 1-161, 'F', 163-661 #label KAW  
KEYWORDS glycoprotein  
SUMMARY #length 661 #molecular-weight 70255 #checksum 5487  
Query Match 100.0%; Score 65; DB 13; Length 661;  
Best Local Similarity 100.0%; Pred. No. 1.13e-03;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 280 ylepdpvta 288  
QY 1 YLEPGPVT A 9  
RESULT 2  
ENTRY A41234 #type complete  
TITLE melanocyte-specific protein Pmel-17 precursor - human

```

ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS    A41234
REFERENCE      A41234
#authors      Kwon, B.S.; Chintanani, C.; Kozak, C.A.; Copeland, N.G.;
              Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
              Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
              coat color locus on mouse chromosome 10 and is in a
              syntenic region on human chromosome 12.
#cross-references MUID:92021023
#accession    A41234
#status       preliminary
#molecule_type mRNA
#residues     1-668 #label KWO
#cross-references GB:W7348
SUMMARY       #length 668 #molecular-weight 70932 #checksum 6409

Query Match      100.0%; Score 65; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 1.13e-03;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 280 ylepgpvt 288
      |||||
QY 1 YLEPGPVT 9

RESULT 3
ENTRY   A49179 #type fragment
TITLE   melanoma antigen homolog rpe1 - bovine (fragment)
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE     19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
23-Mar-1995
ACCESSIONS A49179
REFERENCE  A49179
#authors   Kim, R.Y.; Wistow, G.J.
#journal   Exp. Eye Res. (1992) 55:657-662
#title     The cDNA RPE1 and monoclonal antibody HMB-50 define gene
              products preferentially expressed in retinal pigment
              epithelium.
#cross-references MUID:93122163
#accession A49179
#status     preliminary
#molecule_type nucleic acid
#residues   1-491 #label KIM
#cross-references NCBI:122438; NCBI:122439
#experimental_source retinal pigment epithelium
#note       sequence extracted from NCBI backbone
SUMMARY     #length 491 #checksum 3125

Query Match      89.2%; Score 58; DB 14; Length 491;
Best Local Similarity 88.9%; Pred. No. 4.80e-02;
Matches          8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 123 ylepgpvt 131
      |||||
QY 1 YLEPGPVT 9

RESULT 4
ENTRY   B48205 #type fragment
TITLE   All-1 protein - GTE form - mouse (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
07-Oct-1994
ACCESSIONS B48205
REFERENCE  B48205
#authors   Ma, Q.; Alder, H.; Nelson, K.K.; Chatterjee, D.; Gu, Y.;
              Nakamura, T.; Canaani, E.; Croce, C.M.; Siracusa, L.D.;
              Buchberg, A.M.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1993) 90:6350-6354
#title     Analysis of the murine All-1 gene reveals conserved domains
              with human ALL-1 and identifies a motif shared with DNA
              methyltransferases.
#cross-references GB:L17069
#accession A48205
#status     preliminary
#molecule_type mRNA
#residues   1-3866 #label HAR
#cross-references GB:L17069
SUMMARY     #length 3866 #checksum 2325

Query Match      84.6%; Score 55; DB 14; Length 3866;
Best Local Similarity 77.8%; Pred. No. 2.23e-01;
Matches          7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 2344 ylepgqvtt 2352
      |||||
QY 1 YLEPGPVT 9

RESULT 5
ENTRY   A48205 #type fragment
TITLE   All-1 protein +GTE form - mouse (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
07-Oct-1994
ACCESSIONS A48205
REFERENCE  A48205
#authors   Ma, Q.; Alder, H.; Nelson, K.K.; Chatterjee, D.; Gu, Y.;
              Nakamura, T.; Canaani, E.; Croce, C.M.; Siracusa, L.D.;
              Buchberg, A.M.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1993) 90:6350-6354
#title     Analysis of the murine All-1 gene reveals conserved domains
              with human ALL-1 and identifies a motif shared with DNA
              methyltransferases.
#cross-references GB:L17069
#accession A48205
#status     preliminary
#molecule_type mRNA
#residues   1-3866 #label HAR
#cross-references GB:L17069
SUMMARY     #length 3866 #checksum 2325

Query Match      84.6%; Score 55; DB 14; Length 3866;
Best Local Similarity 77.8%; Pred. No. 2.23e-01;
Matches          7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 2344 ylepgqvtt 2355
      |||||
QY 1 YLEPGPVT 9

RESULT 6
ENTRY   S53871 #type complete
TITLE   Pmel 17 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
03-Nov-1995
ACCESSIONS S53871
REFERENCE  S53871
#authors   Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
              Chintanani, C.; Bennett, D.; Pickard, R.T.
#journal   Nucleic Acids Res. (1995) 23:154-158
#title     Mouse silver mutation is caused by a single base insertion in
              the putative cytoplasmic domain of Pmel 17.
#cross-references GB:L17069
#accession S53871
#status     preliminary
#molecule_type mRNA
#residues   1-626 #label KWO

```

```

#title        Analysis of the murine All-1 gene reveals conserved domains
              with human ALL-1 and identifies a motif shared with DNA
              methyltransferases.
#accession    B48205
#status       preliminary
#molecule_type mRNA
#residues     1-3866 #label HAR
#cross-references GB:L17069
GENETICS
#gene         All-1
KEYWORDS      alternative splicing; zinc finger
SUMMARY       #length 3866 #checksum 2325

Query Match      84.6%; Score 55; DB 14; Length 3866;
Best Local Similarity 77.8%; Pred. No. 2.23e-01;
Matches          7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 2344 ylepgqvtt 2352
      |||||
QY 1 YLEPGPVT 9

RESULT 5
ENTRY   A48205 #type fragment
TITLE   All-1 protein +GTE form - mouse (fragment)
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
07-Oct-1994
ACCESSIONS A48205
REFERENCE  A48205
#authors   Ma, Q.; Alder, H.; Nelson, K.K.; Chatterjee, D.; Gu, Y.;
              Nakamura, T.; Canaani, E.; Croce, C.M.; Siracusa, L.D.;
              Buchberg, A.M.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1993) 90:6350-6354
#title     Analysis of the murine All-1 gene reveals conserved domains
              with human ALL-1 and identifies a motif shared with DNA
              methyltransferases.
#cross-references GB:L17069
#accession A48205
#status     preliminary
#molecule_type mRNA
#residues   1-3866 #label HAR
#cross-references GB:L17069
SUMMARY       #length 3866 #checksum 2325

Query Match      84.6%; Score 55; DB 14; Length 3866;
Best Local Similarity 77.8%; Pred. No. 2.23e-01;
Matches          7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 2344 ylepgqvtt 2355
      |||||
QY 1 YLEPGPVT 9

RESULT 6
ENTRY   S53871 #type complete
TITLE   Pmel 17 protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
03-Nov-1995
ACCESSIONS S53871
REFERENCE  S53871
#authors   Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
              Chintanani, C.; Bennett, D.; Pickard, R.T.
#journal   Nucleic Acids Res. (1995) 23:154-158
#title     Mouse silver mutation is caused by a single base insertion in
              the putative cytoplasmic domain of Pmel 17.
#cross-references GB:L17069
#accession S53871
#status     preliminary
#molecule_type mRNA
#residues   1-626 #label KWO

```

```

SUMMARY          #length 626 #molecular-weight 65979 #checksum 5710
Query Match      78.5%; Score 51; DB 14; Length 626;
Best Local Similarity 77.8%; Pred. No. 1.61e+00;
Matches          7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 280 ylesgsyta 288
|||||
QY 1 YLEPGPVTA 9

RESULT 7
ENTRY   A44264 #type fragment
TITLE   trithorax homolog HTX (version 1) - human (fragment)
ALTERNATE_NAMES trithorax homolog ALL-1
ORGANISM #formal_name Homo sapiens #common_name man
DATE     10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change
        06-Sep-1996
ACCESSIONS A44264
REFERENCE  A44264
#authors   Gu, Y.; Nakamura, T.; Alder, H.; Prasad, R.; Canaanl, O.;
#journal   Cimino, G.; Croce, C.M.; Canaanl, E.
#title     Cell (1992) 71:701-708
#note      The t(4;11) chromosome translocation of human acute leukemias
           fuses the ALL-1 gene, related to Drosophila trithorax, to
           the AP-4 gene.
#cross-references MUID:93046668
#accession A44264
#status    not compared with conceptual translation
#molecule_type mRNA
#residues  1-3910 #label GUL
#cross-references NCBI:11779
#note      sequence extracted from NCBI backbone

GENETICS
#gene      GDB:MLL; HTX; ALL-1
#cross-references GDB:128819
#map_position llq23-llq23
KEYWORDS   proto-oncogene; zinc finger
SUMMARY    #length 3910 #checksum 5875

Query Match      76.9%; Score 50; DB 13; Length 3910;
Best Local Similarity 66.7%; Pred. No. 2.61e+00;
Matches          6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 2387 flepgqvtt 2395
|||||
QY 1 YLEPGPVTA 9

RESULT 8
ENTRY   A44265 #type complete
TITLE   trithorax homolog HTX (version 2) - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE     30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change
        06-Sep-1996
ACCESSIONS A44265
REFERENCE  A44265
#authors   Tkachuk, D.C.; Kohler, S.; Cleary, M.L.
#journal   Cell (1992) 71:691-700
#title     Involvement of a homolog of Drosophila trithorax by llq23
           chromosomal translocations in acute leukemias.
#cross-references MUID:93046667
#accession A44265
#status    nucleic acid sequence not shown; not compared with
           conceptual translation
#molecule_type mRNA
#residues  1-3968 #label TKA
#cross-references NCBI:11779
#note      sequence extracted from NCBI backbone

GENETICS
#gene      GDB:MLL; HTX; ALL-1
#cross-references GDB:128819
#map_position llq23-llq23

```

---

```

KEYWORDS         proto-oncogene; zinc finger
SUMMARY          #length 3968 #molecular-weight 431942 #checksum 8269

Query Match      76.9%; Score 50; DB 13; Length 3968;
Best Local Similarity 66.7%; Pred. No. 2.61e+00;
Matches          6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 2446 flepgqvtt 2454
|||||
QY 1 YLEPGPVTA 9

RESULT 9
ENTRY   D49599 #type complete
TITLE   matrix protein M2 - hemorrhagic septicemia virus (strain
        Makah)
ORGANISM #formal_name hemorrhagic septicemia virus
DATE     01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change
        01-Dec-1995
ACCESSIONS D49599
REFERENCE  A49599
#authors   Benmansour, A.; Paubert, G.; Bernard, J.; De Kinkelin, P.
#journal   Virology (1994) 198:602-612
#title     The polymerase-associated protein (M1) and the matrix protein
           (M2) from a virulent and an avirulent strain of viral
           hemorrhagic septicemia virus (VHSV), a fish rhabdovirus.
#accession D49599
#status    preliminary
#molecule_type mRNA
#residues  1-201 #label BEN
#cross-references GB:003503
SUMMARY    #length 201 #molecular-weight 22262 #checksum 995

Query Match      75.4%; Score 49; DB 8; Length 201;
Best Local Similarity 71.4%; Pred. No. 4.19e+00;
Matches          5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 109 yldpspv 115
|||||
QY 1 YLEPGPV 7

RESULT 10
ENTRY   S01929 #type complete
TITLE   fimbrial protein precursor - Bordetella pertussis
ORGANISM #formal_name Bordetella pertussis
DATE     18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change
        03-May-1996
ACCESSIONS S01929
REFERENCE  S01929
#authors   Pedroni, P.; Riboli, B.; de Ferra, F.; Grandi, G.; Toma, S.;
           Arico, B.; Rappuoli, R.
#journal   Mol. Microbiol. (1988) 2:539-543
#title     Cloning of a novel pilin-like gene from Bordetella pertussis:
           homology to the fim2 gene.
#cross-references MUID:89013896
#accession S01929
#molecule_type DNA
#residues  1-201 #label PED
#cross-references EMBL:Y00556

GENETICS
#gene      fimX
CLASSIFICATION #superfamily type 1 fimbrial protein
FEATURE
1-21         #domain signal sequence #label SIG\
22-201       #product fimbrial protein #label MAT
SUMMARY      #length 201 #molecular-weight 21447 #checksum 5799

Query Match      72.3%; Score 47; DB 7; Length 201;
Best Local Similarity 75.0%; Pred. No. 1.06e+01;
Matches          6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 89 yfepgptt 96

```



```

QY      1 YLEPGPVT 8
      1:||||| I
RESULT  11
ENTRY   S36451      #type complete
TITLE   fimbrial protein fimX - Bordetella bronchiseptica
ORGANISM #formal_name Bordetella bronchiseptica
DATE    06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
ACCESSIONS S36451
REFERENCE  S36451
#authors   Savelkoul, P.; de Kerf, D.; de Groot, L.; Willems, R.; Mooi,
           F.; van der Zeijst, B.; Gastra, W.
#submission Submitted to the EMBL Data Library, February 1993
#description Characterization of fimX, a gene encoding a fimbrial subunit
           protein of Bordetella bronchiseptica.
#accession S36451
##status   preliminary
##molecule_type DNA
##residues 1-201 #label SAV
##cross-references EMBL:X74118
GENETICS
#gene      fimX
CLASSIFICATION #superfamily type 1 fimbrial protein
SUMMARY      #length 201 #molecular-weight 21462 #checksum 5181
Query Match  72.3%; Score 47; DB 7; Length 201;
Best Local Similarity 75.0%; Pred. No. 1.06e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 89 yfepgptt 96
      1:||||| I
QY      1 YLEPGPVT 8

RESULT  12
ENTRY   S03754      #type complete
TITLE   fimbrial protein precursor - Bordetella pertussis (serotype
           2, strain Wellcome 28)
ORGANISM #formal_name Bordetella pertussis
DATE    31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change
ACCESSIONS S03754
REFERENCE  S03754
#authors   Livey, I.; Duggleby, C.J.; Robinson, A.
#journal   Mol. Microbiol. (1987) 1:203-209
#title     Cloning and nucleotide sequence analysis of the serotype 2
           fimbrial subunit gene of Bordetella pertussis.
#cross-references MUID:88216164
#accession S03754
##molecule_type DNA
##residues 1-207 #label LIV
##cross-references EMBL:Y00527
##note     part of this sequence, including the amino end of the
           mature protein, was confirmed by protein sequencing
CLASSIFICATION #superfamily type 1 fimbrial protein
FEATURE 1-26      #domain signal sequence #status experimental #label SIG\
27-207      #product fimbrial protein #status experimental #label
           MAT
SUMMARY      #length 207 #molecular-weight 21950 #checksum 9877
Query Match  72.3%; Score 47; DB 7; Length 207;
Best Local Similarity 75.0%; Pred. No. 1.06e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 96 yfepgptt 103
      1:||||| I
QY      1 YLEPGPVT 8

RESULT  13
ENTRY   A55347      #type complete
TITLE   adhesin ligand epiligrin, alpha-3 chain form A precursor -
           human
ORGANISM #formal_name Homo sapiens #common_name man
DATE    06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change
ACCESSIONS A55347
REFERENCE  A55347
#authors   Ryan, M.C.; Tizard, R.; VanDevanter, D.R.; Carter, W.G.

```

---

```

ENTRY   S36449      #type complete
TITLE   fimbrial protein fim2 - Bordetella bronchiseptica
ORGANISM #formal_name Bordetella bronchiseptica
DATE    06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
ACCESSIONS S36449
REFERENCE  S36449
#authors   Savelkoul, P.; de Kerf, D.; Mooi, F.; van der Zeijst, B.;
           Gastra, W.
#submission Submitted to the EMBL Data Library, February 1993
#description Characterization of the fim2 and fim3 fimbrial subunit genes
           of Bordetella bronchiseptica; role of Fim2 and Fim3
           fimbriae and flagella in adhesion.
#accession S36449
##status   preliminary
##molecule_type DNA
##residues 1-208 #label SAV
##cross-references EMBL:X74119
GENETICS
#gene      fim2
CLASSIFICATION #superfamily type 1 fimbrial protein
SUMMARY      #length 208 #molecular-weight 21837 #checksum 536
Query Match  72.3%; Score 47; DB 7; Length 208;
Best Local Similarity 75.0%; Pred. No. 1.06e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 96 yfepgptt 103
      1:||||| I
QY      1 YLEPGPVT 8

RESULT  14
ENTRY   B36186      #type complete
TITLE   I factor 2 (transposon) - fruit fly (Drosophila teissieri)
ORGANISM #formal_name Drosophila teissieri
DATE    14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change
ACCESSIONS B36186
REFERENCE  B36186
#authors   Abad, P.; Vaury, C.; Pelisson, A.; Chaboissier, M.C.;
           Buseau, I.; Bucheton, A.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1989) 86:8887-8891
#title     A long interspersed repetitive element--the I factor of
           Drosophila teissieri--is able to transpose in different
           Drosophila species.
#cross-references MUID:90046894
#accession B36186
##status   preliminary
##molecule_type DNA
##residues 1-1234 #label ABA
##cross-references GB:M28878
SUMMARY      #length 1234 #molecular-weight 141025 #checksum 1614
Query Match  72.3%; Score 47; DB 12; Length 1234;
Best Local Similarity 62.5%; Pred. No. 1.06e+01;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1148 ylnpnpit 1155
      1:||||| I
QY      1 YLEPGPVT 8

RESULT  15
ENTRY   A55347      #type complete
TITLE   adhesin ligand epiligrin, alpha-3 chain form A precursor -
           human
ORGANISM #formal_name Homo sapiens #common_name man
DATE    06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change
ACCESSIONS A55347
REFERENCE  A55347
#authors   Ryan, M.C.; Tizard, R.; VanDevanter, D.R.; Carter, W.G.

```

```
#journal      J. Biol. Chem. (1994) 269:22779-22787
#title        Cloning of the lamA3 gene encoding the alpha3 chain of the
#             adhesive ligand epiligrin. Expression in wound repair.
#accession    A55347
#status       preliminary
#molecule_type mRNA
#residues     1-1713 #label RYA
#cross-references GB:L34135
KEYWORDS      alternative splicing
SUMMARY       #length 1713 #molecular-weight 189304 #checksum 8941

Query Match      72.3%; Score 47; DB 13; Length 1713;
Best Local Similarity 77.8%; Pred. No. 1.06e+01;
Matches          7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1589 yleagkvta 1597
   |||:| |||
QY  1 YLEPGPVT 9

Search completed: Tue Jun 10 11:31:27 1997
Job time : 10 secs.
```

\*\*\*\*\*

WQEQELH (TM)

\*\*\*\*\*

Release 2.1d John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:30:50 1997; MasPar time 2.01 Seconds  
Tabular output not generated. 95.098 Million cell updates/sec

Title: >US-08-231-565A-38  
Description: (1-9) from US08231565A.pap  
Perfect score: 65  
Sequence: 1 YLEPQPVTA 9

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 23.012; Variance 23.930; scale 0.962

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	65	100.0	668	PM17_HUMAN	PMEL 17 PROTEIN PRECU	7.32e-05
2	55	84.6	3866	HRX_MOUSE	ZINC FINGER PROTEIN H	3.52e-02
3	50	76.9	3969	HRX_HUMAN	ZINC FINGER PROTEIN H	6.13e-01
4	47	72.3	201	FMFX_BORPE	FIMBRIAL PROTEIN FMX	3.11e+00
5	47	72.3	207	FM2_BORPE	SEROTYPE 2 FIMBRIAL S	3.11e+00
6	47	72.3	2257	DPOE_HUMAN	DNA POLYMERASE EPSILO	3.11e+00
7	46	70.8	747	AMD1_RAT	AMP DEAMINASE 1 (EC 3	5.25e+00
8	46	70.8	1010	CONT_CHICK	CONTACTIN PRECURSOR (	5.25e+00
9	45	69.2	346	YOJK_HAEIN	HYPOTHETICAL PROTEIN	8.79e+00
10	45	69.2	501	VLL_PAPVE	MAJOR CAPSID PROTEIN	8.79e+00
11	45	69.2	513	VLL_PAPVD	MAJOR CAPSID PROTEIN	8.79e+00
12	45	69.2	555	NRM1_CHICK	NATURAL RESISTANCE-AS	8.79e+00
13	44	67.7	152	NKDE_HUMAN	NUCLEOSIDE DIPHOSPHAT	1.46e+01
14	44	67.7	152	NKDB_MOUSE	NUCLEOSIDE DIPHOSPHAT	1.46e+01
15	44	67.7	152	NKDB_RAT	NUCLEOSIDE DIPHOSPHAT	1.46e+01
16	44	67.7	504	VLL_HPV51	MAJOR CAPSID PROTEIN	1.46e+01
17	44	67.7	602	YH02_YEAST	HYPOTHETICAL 69.0 KD	1.46e+01
18	43	66.2	152	NKDE_RAT	NUCLEOSIDE DIPHOSPHAT	2.39e+01
19	43	66.2	152	NKDE_MOUSE	NUCLEOSIDE DIPHOSPHAT	2.39e+01
20	43	66.2	152	NKDA_BOVIN	NUCLEOSIDE DIPHOSPHAT	2.39e+01
21	43	66.2	152	NKDA_HUMAN	NUCLEOSIDE DIPHOSPHAT	2.39e+01
22	43	66.2	152	NKDE_BOVIN	NUCLEOSIDE DIPHOSPHAT	2.39e+01

23	43	66.2	351	10	WNT4_MOUSE	WNT-4 PROTEIN PRECURS	2.39e+01
24	43	66.2	443	8	PYRO_SCHPO	DIHYDROROTATE DEHYDR	2.39e+01
25	43	66.2	468	10	WNTG_DROME	PROTEIN DINT-1 PRECUR	2.39e+01
26	43	66.2	481	5	KRL1_HSV2	SERINE/THREONINE-PROT	2.39e+01
27	43	66.2	486	6	MVL_DROME	MALVOLIO PROTEIN.	2.39e+01
28	43	66.2	504	6	NRM2_HUMAN	NATURAL RESISTANCE-AS	2.39e+01
29	43	66.2	505	10	VLL_CRPVK	MAJOR CAPSID PROTEIN	2.39e+01
30	43	66.2	513	7	PEPB_SOLTU	PYROPHOSPHATE--FRUCTO	2.39e+01
31	43	66.2	514	6	NRM2_MOUSE	NATURAL RESISTANCE-AS	2.39e+01
32	43	66.2	914	4	GUX2_CLOSR	EXOGLUCANASE II PRECU	2.39e+01
33	43	66.2	1040	8	RAG1_MOUSE	V(D)J RECOMBINATION A	2.39e+01
34	43	66.2	3027	7	POLG_PIFV1	GENOME POLYPROTEIN (C	2.39e+01
35	43	66.2	3898	7	POLG_HCVA	GENOME POLYPROTEIN.	2.39e+01
36	43	66.2	3898	7	POLG_BVDVS	GENOME POLYPROTEIN.	2.39e+01
37	43	66.2	3988	7	POLG_BVDVN	GENOME POLYPROTEIN.	2.39e+01
38	42	64.6	151	6	NDK_GINCI	NUCLEOSIDE DIPHOSPHAT	3.89e+01
39	42	64.6	499	10	VLL_HPV53	MAJOR CAPSID PROTEIN	3.89e+01
40	42	64.6	502	10	VLL_HPV28	MAJOR CAPSID PROTEIN	3.89e+01
41	42	64.6	510	10	VLL_HPV57	MAJOR CAPSID PROTEIN	3.89e+01
42	42	64.6	510	10	VLL_HPV2A	MAJOR CAPSID PROTEIN	3.89e+01
43	42	64.6	534	10	VLL_HPV56	MAJOR CAPSID PROTEIN	3.89e+01
44	42	64.6	895	4	GNDS_RAT	GUANINE NUCLEOTIDE DI	3.89e+01
45	42	64.6	2142	1	BAT2_HUMAN	LARGE PROLINE-RICH PR	3.89e+01

ALIGNMENTS

RESULT 1  
ID PM17\_HUMAN STANDARD; PRT; 668 AA.  
AC P40967;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE PMEL 17 PROTEIN PRECURSOR.  
GN PMEL17.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92021023.  
RA KWON B.S., CHINTAMANANI C., KOZAK C.A., COPELAND N.G.,  
RA GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,  
RA KIM K.-K.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).  
CC -!- FUNCTION: COULD BE A MELANOGENIC ENZYME.  
CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).  
DR EMBL; M77348; G190106; -.  
DR MIM; 155550; -.  
KW TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 668 PMEL 17 PROTEIN.  
FT TRANSMEM 575 595 POTENTIAL.  
FT TRANSMEM 603 623 POTENTIAL.  
FT DOMAIN 217 307 PRD.  
FT DOMAIN 315 444 10 X 13 AA TANDEM REPEATS.  
FT REPEAT 328 340 2.  
FT REPEAT 341 353 3.  
FT REPEAT 354 366 4.  
FT REPEAT 367 379 5.  
FT REPEAT 380 392 6.  
FT REPEAT 393 405 7.  
FT REPEAT 406 418 8.  
FT REPEAT 419 431 9.  
FT REPEAT 432 444 10.  
FT CARBOHYD 81 81 POTENTIAL.  
FT CARBOHYD 106 106 POTENTIAL.  
FT CARBOHYD 111 111 POTENTIAL.  
FT CARBOHYD 321 321 POTENTIAL.  
FT CARBOHYD 568 568 POTENTIAL.  
SQ SEQUENCE 668 AA; 70992 MW; 6E8E1AF0 CRC32;

```

Query Match      100.0%; Score 65; DB 7; Length 668;
Best Local Similarity 100.0%; Pred. No. 7.32e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 280 ylepqpvt 288
      |||||
QY 1 YLEPGPVT 9

RESULT 2
ID HRX_MOUSE STANDARD; PRT; 3866 AA.
AC P55200;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE ZINC FINGER PROTEIN HRX (ALL-1) (FRAGMENT).
GN MLL OR HRX OR ALL1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J AND B6/CBA; TISSUE=SPLEEN, AND LUNG;
RX MEDLINE; 93317679.
RA MA Q., ALDER H., NELSON K.K., CHATTERJEE D., GU Y., NAKAMURA T.,
RA CANAANI E., CROCE C.M., SIRACUSA L.D., BUCHBERG A.M.;
RL PROC. NATL. ACAD. SCI. U.S.A. 90:6350-6354(1993).
CC -1- FUNCTION: POSSIBLY ACTS AS A TRANSCRIPTIONAL REGULATORY FACTOR.
CC MAY REGULATE GENES INVOLVED IN SKELETAL FORMATION DURING
CC EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: TO D.MELANOGASTER TRITHORAX.
CC -1- SIMILARITY: CONTAINS A COPY OF THE BROMODOMAIN.
CC -1- SIMILARITY: CONTAINS A 'SET' DOMAIN.
DR EMBL; L17069; G688443; -.
KW PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION; DNA-BINDING;
KW NUCLEAR PROTEIN; ZINC-FINGER; METAL-BINDING; TRANSCRIPTION REGULATION;
KW ALTERNATIVE SPLICING; POLYMORPHISM.
FT NON_TER 1 1
FT DOMAIN 1 118 ARG/CYS/LYS-RICH (DNA MTASE-LIKE).
FT DNA_BIND 35 41 POLY-GLY.
FT DNA_BIND 67 78 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 115 125 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 199 207 A.T HOOK (BY SIMILARITY).
FT DOMAIN 459 469 PRO-RICH.
FT DOMAIN 1047 1118 ARG/CYS/LYS-RICH (DNA MTASE-LIKE).
FT DOMAIN 1231 1238 POLY-PRO.
FT ZN_FING 1333 1351 BY SIMILARITY.
FT ZN_FING 1355 1378 BY SIMILARITY.
FT ZN_FING 1381 1400 BY SIMILARITY.
FT ZN_FING 1405 1429 BY SIMILARITY.
FT ZN_FING 1432 1458 BY SIMILARITY.
FT ZN_FING 1468 1490 BY SIMILARITY.
FT ZN_FING 1495 1526 BY SIMILARITY.
FT ZN_FING 1605 1650 BROMODOMAIN (DIVERGENT).
FT ZN_FING 1775 1805 BY SIMILARITY.
FT ZN_FING 1835 1852 BY SIMILARITY.
FT ZN_FING 1857 1879 BY SIMILARITY.
FT DOMAIN 2163 2166 POLY-SER.
FT DOMAIN 2765 2768 POLY-SER.
FT DOMAIN 3533 3536 POLY-GLU.
FT DOMAIN 3693 3697 POLY-GLU.
FT DOMAIN 3737 3866 SET DOMAIN.
FT VARIANT 1497 1497 K -> T.
FT VARSPLIC 1503 1505 MISSING (IN ALTERNATIVE SPLICED FORM).
SQ SEQUENCE 3866 AA; 420976 MW; 200FC2A1 CRC32;

Query Match      84.6%; Score 55; DB 5; Length 3866;
Best Local Similarity 77.8%; Pred. No. 3.52e-02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 2344 ylepqpvt 2352
      |||||

```

```

QY 1 YLEPGPVT 9
RESULT 3
ID HRX_HUMAN STANDARD; PRT; 3969 AA.
AC Q03164;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ZINC FINGER PROTEIN HRX (ALL-1).
GN MLL OR HRX OR ALL1 OR TRX1 OR HTRX.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93046667.
RA TRACHUK D.C., KOHLER S., CLEARY M.L.;
RN CELL 71:691-700(1992).
RN [2]
RP SEQUENCE OF 1317-2328 FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE; 93265134.
RA DJABALI M., SELLERI L., PARRY P., BOWER M., YOUNG B.D.,
RA EVANS G.A.;
RN NAT. GENET. 2:113-118(1992).
RN [3]
RP SEQUENCE OF 1251-1538 FROM N.A.
RA GU Y., ALDER H., NAKAMURA T., SCHICHMAN S.A., PRASAD R.,
RA CANAANI O., SAITO H., CROCE C.M., CANAANI E.;
RN CANCER RES. 54:2327-2330(1994).
RN [4]
RP IDENTIFICATION OF PROBABLE FRAMESHIFT IN 317-379.
RA ASLAND R.;
RP UNPUBLISHED OBSERVATIONS (MAR-1995).
CC -1- FUNCTION: POSSIBLY ACTS AS A TRANSCRIPTIONAL REGULATORY FACTOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: HEART, LUNG, BRAIN AND T AND B LYMPHOCYTES.
CC -1- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY CHROMOSOMAL TRANSLOCATIONS
CC T(11;19)(Q23;P13.3) THAT INVOLVES MLL AND ENL; T(4;11)(Q21;Q23)
CC THAT INVOLVES MLL AND AF4; T(9;11)(P22;Q23) THAT INVOLVES MLL AND
CC AF9; T(6;11)(Q27;Q23) THAT INVOLVES MLL AND AF6; T(X;11)(Q13;Q23)
CC THAT INVOLVES MLL AND AFX; T(11;17)(Q23;Q21) THAT INVOLVES MLL AND
CC AF17; T(10;11)(P12;Q23) THAT INVOLVES MLL AND AF10 AND
CC T(11;19)(Q23;P13.3) THAT INVOLVES MLL AND ELL.
CC -1- SIMILARITY: TO D.MELANOGASTER TRITHORAX.
CC -1- SIMILARITY: CONTAINS A COPY OF THE BROMODOMAIN.
CC -1- SIMILARITY: CONTAINS A 'SET' DOMAIN.
DR EMBL; L04284; G184394; -.
DR EMBL; L01986; G553800; -.
DR EMBL; U04737; G451555; -.
DR HSSP; P19999; ICLG.
DR MIN; 159555; -.
DR PROSITE; PS50014; BROMODOMAIN_2.
KW PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION; DNA-BINDING;
KW NUCLEAR PROTEIN; ZINC-FINGER; METAL-BINDING; TRANSCRIPTION REGULATION.
FT DOMAIN 17 102 ALA/GLY/SER-RICH.
FT DOMAIN 17 23 POLY-GLY.
FT DOMAIN 61 67 POLY-ALA.
FT DOMAIN 85 101 POLY-SER.
FT DNA_BIND 169 180 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 217 227 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 301 309 A.T HOOK (BY SIMILARITY).
FT ZN_FING 1434 1452 BY SIMILARITY.
FT ZN_FING 1456 1479 BY SIMILARITY.
FT ZN_FING 1482 1501 BY SIMILARITY.
FT ZN_FING 1506 1530 BY SIMILARITY.
FT ZN_FING 1533 1559 BY SIMILARITY.
FT ZN_FING 1569 1591 BY SIMILARITY.
FT ZN_FING 1596 1624 BY SIMILARITY.
FT DOMAIN 1703 1748 BROMODOMAIN (DIVERGENT).
FT ZN_FING 1873 1903 BY SIMILARITY.
FT ZN_FING 1933 1950 BY SIMILARITY.

```

```

FT ZN_FING 1955 1977 BY SIMILARITY.
FT DOMAIN 3840 3969 SET DOMAIN.
FT CONFLICT 317 379 GLLINSELEKPKQVRKDEGTPLTKEDKTVVQSPRRIKP
FT KTRIPSSKRTDIAIKQLQRA -> VSSLILNWKSPKSG
FT KTRKEHLQKRIQLSKALEGLSOLGLFLQKGMQPLL
FT SNSYRGO (IN REF. 1).
FT R -> G (IN REF. 3).
FT CONFLICT 1487 1487
FT CONFLICT 1937 1937 Q -> H (IN REF. 2).
FT CONFLICT 2181 2181 P -> S (IN REF. 2).
SQ SEQUENCE 3969 AA; 431882 MW; 93D16981 CRC32;

Query Match 76.9%; Score 50; DB 5; Length 3969;
Best Local Similarity 66.7%; Pred. No. 6.13e-01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 2446 flepqgvt 2454
:|||||
QY 1 YLEPGPVT 9

RESULT 4
ID FMFX BORPE STANDARD; PRT; 201 AA.
AC P09808;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE FIMBRIAL PROTEIN FIMX PRECURSOR (PILIN).
GN FIMX.
OS BORDETELLA PERTUSSIS
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC ALCALIGENACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-SAL;
RX MEDLINE; 89013896.
RA PEDRONI P., RIBOLI B., DE FERRA F., GRANDI G., TOMA S., ARICO B.,
RA RAPPUOLI R.;
RL MOL. MICROBIOL. 2:539-543(1988).
CC -!- DISEASE: BORDETELLA PERTUSSIS IS THE CAUSATIVE AGENT OF WHOOPING
CC COUGH. AN ESSENTIAL STEP IN THE DISEASE PROCESS IS THE ATTACHMENT
CC OF THE BACTERIA TO THE CILIATED EPITHELIUM OF THE RESPIRATORY
CC TRACT, ENABLING THE ORGANISM TO RESIST NORMAL HOST-CLEARANCE
CC MECHANISMS. IT IS UNCLEAR WHICH BACTERIAL CELL SURFACE COMPONENT
CC ARE RESPONSIBLE FOR ADHERENCE BUT THE FIMBRIAE OF B.PERTUSSIS ARE
CC PRIME CANDIDATES FOR BEING INVOLVED IN THIS PROCESS.
CC -!- SUBCELLULAR LOCATION: PILI STRUCTURE ON THE CELL SURFACE.
CC -!- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.
DR EMBL; Y00556; G39751; -.
DR EMBL; A25464; G904626; -.
DR PIR; S01929; S01929.
KW FIMBRIA; SIGNAL.
FT SIGNAL 1 21
FT CHAIN 22 201 FIMBRIAL PROTEIN FIMX.
FT DISULFID 37 79 PROBABLE.
SQ SEQUENCE 201 AA; 21447 MW; 94C8DD3D CRC32;

Query Match 72.3%; Score 47; DB 4; Length 201;
Best Local Similarity 75.0%; Pred. No. 3.11e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 89 yfepgptt 96
:|||||
QY 1 YLEPGPVT 8

RESULT 5
ID FM2 BORPE STANDARD; PRT; 207 AA.
AC P05788;
DT 01-NOV-1988 (REL. 09, CREATED)
DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE SEROTYPE 2 FIMBRIAL SUBUNIT PRECURSOR.
GN FIM2.

```

```

OS BORDETELLA PERTUSSIS.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC ALCALIGENACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=WELLCOME 28;
RX MEDLINE; 89216164;
RA LIVEX I., DUGGLEBY C.J., ROBINSON A.;
RL MOL. MICROBIOL. 1:203-209(1987).
CC -!- DISEASE: BORDETELLA PERTUSSIS IS THE CAUSATIVE AGENT OF WHOOPING
CC COUGH. AN ESSENTIAL STEP IN THE DISEASE PROCESS IS THE ATTACHMENT
CC OF THE BACTERIA TO THE CILIATED EPITHELIUM OF THE RESPIRATORY
CC TRACT, ENABLING THE ORGANISM TO RESIST NORMAL HOST-CLEARANCE
CC MECHANISMS. IT IS UNCLEAR WHICH BACTERIAL CELL SURFACE COMPONENT
CC ARE RESPONSIBLE FOR ADHERENCE BUT THE FIMBRIAE OF B.PERTUSSIS ARE
CC PRIME CANDIDATES FOR BEING INVOLVED IN THIS PROCESS.
CC -!- SUBCELLULAR LOCATION: PILI STRUCTURE ON THE CELL SURFACE.
CC -!- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.
DR EMBL; Y00527; G40343; -.
DR PIR; S03754; S03754.
KW FIMBRIA; SIGNAL.
FT SIGNAL 1 26
FT CHAIN 27 207 SEROTYPE 2 FIMBRIAL SUBUNIT.
FT DISULFID 42 85 PROBABLE.
SQ SEQUENCE 207 AA; 21950 MW; EF348AC2 CRC32;

Query Match 72.3%; Score 47; DB 4; Length 207;
Best Local Similarity 75.0%; Pred. No. 3.11e+00;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 96 yfepgptt 103
:|||||
QY 1 YLEPGPVT 8

RESULT 6
ID DPOE_HUMAN STANDARD; PRT; 2257 AA.
AC Q07864;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE DNA POLYMERASE EPSILON, CATALYTIC SUBUNIT A (EC 2.7.7.7) (DNA
DE POLYMERASE II SUBUNIT A).
GN POLE.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 48-51; 876-886 AND 1338-1344.
RX MEDLINE; 93252906.
RA KESTI T., FRANTTI H., SYVAOJA J.E.;
RL J. BIOL. CHEM. 268:10238-10245(1993).
CC -!- FUNCTION: PARTICIPATES IN DNA REPAIR AND IN CHROMOSOMAL DNA
CC REPLICATION.
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC N PYROPHOSPHATE + DNA(N).
CC -!- SUBUNIT: CONSISTS OF TWO SUBUNITS (258 KD AND 55 KD).
CC -!- DOMAIN: THE DNA POLYMERASE ACTIVITY DOMAIN RESIDES IN THE
CC N-TERMINAL HALF OF THE PROTEIN, WHILE THE C-TERMINUS IS NECESSARY
CC FOR COMPLEXING SUBUNIT B AND C. THE C-TERMINUS MAY ALSO REGULATE
CC THE CATALYTIC ACTIVITIES OF THE ENZYME.
CC -!- IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES: ALPHA, BETA, GAMMA,
CC DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR DIFFERENT REACTIONS
CC OF DNA SYNTHESIS.
CC -!- SIMILARITY: BELONGS TO FAMILY B OF DNA POLYMERASES. HIGH
CC SIMILARITY WITH YEAST DNA POLYMERASE II.
DR EMBL; S60080; G303157; -.
DR EMBL; L09561; G915268; ALT_SEQ.
DR PIR; A46692; A46692.
DR MM; 174762; -.
KW DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION; DNA REPAIR; DNA-BINDING;
KW ZINC-FINGER.
GN FIM2.

```

SQ SEQUENCE 2257 AA; 257967 MW; F0166BF2 CRC32;  
 Query Match 72.3%; Score 47; DB 3; Length 2257;  
 Best Local Similarity 71.4%; Pred. No. 3.11e+00;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 1449 ylepgsi 1455  
 |||||  
 QY 1 YLEPGPV 7

RESULT 7  
 ID AMD1\_RAT STANDARD; PRT; 747 AA.  
 AC P10759;  
 DT 01-JUL-1989 (REL. 11, CREATED)  
 DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE AMP DEAMINASE 1 (EC 3.5.4.6) (MYOADENYLATE DEAMINASE) (AMP DEAMINASE ISOFORM M).  
 DE ISOFORM M).  
 GN AMPD1.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 [1]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 536-548.  
 RP TISSUE=MUSCLE;  
 RC MEDLINE; 87308255.  
 RX SABINA R.L., MARQUETANT R., DESAI N.M., KALETHA K., HOLMES E.W.;  
 RA J. BIOL. CHEM. 262:12397-12400(1987).  
 CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY METABOLISM.  
 CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
 CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE, EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3 IS FOUND IN ERYTHROCYTES.  
 CC -1- SIMILARITY: TO AMP DEAMINASE FROM OTHERS SPECIES AND TO ADENOSINE DEAMINASE.  
 DR EMBL; J02811; G202883; -.  
 DR PIR; A27366; A27366.  
 DR PROSITE; PS00485; A\_DEAMINASE.  
 KW HYDROLASE; NUCLEOTIDE METABOLISM; MULTIGENE FAMILY.  
 FT ACT\_SITE 363 363 POTENTIAL.  
 FT ACT\_SITE 573 573 POTENTIAL.  
 FT ACT\_SITE 649 649 POTENTIAL.  
 FT ACT\_SITE 650 650 POTENTIAL.  
 SQ SEQUENCE 747 AA; 86431 MW; 47318096 CRC32;

Query Match 70.8%; Score 46; DB 1; Length 747;  
 Best Local Similarity 85.7%; Pred. No. 5.25e+00;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 705 ylepgpv 711  
 |||||  
 QY 1 YLEPGPV 7

RESULT 8  
 ID CONT\_CHICK STANDARD; PRT; 1010 AA.  
 AC P14781; P10450;  
 DT 01-MAR-1989 (REL. 10, CREATED)  
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
 DE CONTACTIN PRECURSOR (NEURAL CELL RECOGNITION MOLECULE FII).  
 OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;  
 OC GALLIFORMES.  
 GN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE; 90180453.  
 RX

RA BRUEMMENDORF T., WOLFF J.M., RAINER F., RATHJEN F.G.;  
 RL NEURON 2:1351-1361(1989).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WHITE LEGHORN;  
 RX MEDLINE; 89008597.  
 RA RANSCHT B., DOURS M.T.;  
 RL J. CELL BIOL. 107:1561-1573(1988).  
 [3]  
 RN GPI-ANCHOR.  
 RP GPI-ANCHOR.  
 RX MEDLINE; 89286606.  
 RA WOLFF J.M., BRUEMMENDORF T., RATHJEN F.G.;  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 161:931-938(1989).  
 CC -1- FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS SYSTEM DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS SIX C2-LIKE DOMAINS FOLLOWED BY FOUR FIBRONECTIN TYPE III-LIKE DOMAINS.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN IN THE C-TERMINUS AND IS LONGER DUE TO A FRAMESHIFT.  
 CC EMBL; X14877; G63385; -.  
 DR EMBL; Y00813; G63329; ALT\_FRAME.  
 DR PIR; JU0094; JU0094.  
 DR PIR; S01998; S01998.  
 DR HSSP; P20241; 1CFB.  
 KW IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; GPI-ANCHOR; CELL ADHESION;  
 KW REPEAT.  
 FT SIGNAL 1 19  
 FT CHAIN 20 ?  
 FT PROPEP 2 1010  
 FT DOMAIN 50 113  
 FT DOMAIN 143 210  
 FT DOMAIN 247 308  
 FT DOMAIN 336 389  
 FT DOMAIN 420 482  
 FT DOMAIN 510 581  
 FT DOMAIN 593 599  
 FT DOMAIN 600 701  
 FT DOMAIN 702 803  
 FT DOMAIN 804 900  
 FT DOMAIN 901 996  
 FT CARBOHYD 200 200  
 FT CARBOHYD 249 249  
 FT CARBOHYD 329 329  
 FT CARBOHYD 448 448  
 FT CARBOHYD 464 464  
 FT CARBOHYD 485 485  
 FT CARBOHYD 512 512  
 FT CARBOHYD 582 582  
 FT CARBOHYD 924 924  
 SQ SEQUENCE 1010 AA; 112507 MW; 4B397AE4 CRC32;

Query Match 70.8%; Score 46; DB 2; Length 1010;  
 Best Local Similarity 77.8%; Pred. No. 5.25e+00;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 269 ylempata 277  
 |||||  
 QY 1 YLEPGPVTA 9

RESULT 9  
 ID YOJK\_HAEIN STANDARD; PRT; 346 AA.  
 AC P44550;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL PROTEIN HI0172.  
 GN HI0172.  
 OS HAEMOPHILUS INFLUENZAE.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC PASTEURRELLACEAE.

```

RN  SEQUENCE FROM N.A.
RP  STRAIN-RD / KW20;
RX  MEDLINE; 95350630.
RA  FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA  KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA  KRENNENY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA  SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA  WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA  UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA  FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA  GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA  VENTER J.C.;
RL  SCIENCE 269:496-512(1995).
RC  -!- SIMILARITY: STRONG, TO E.COLI YOJK.
DR  EMBL; I44817; G1003244; -.
DR  EMBL; U32702; G925399; -.
KW  HYPOTHETICAL PROTEIN.
SQ  SEQUENCE 346 AA; 37780 MW; 44BF78C3 CRC32;

Query Match      69.2%; Score 45; DB 11; Length 346;
Best Local Similarity 55.6%; Pred. No. 8.79e+00;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db  42 yldgsita 50
    ||: ||: ||
QY  1 YLEPGPVTA 9

RESULT 10
ID  VLI_PAPVE      STANDARD;      PRT;    501 AA.
AC  P11326;
DT  01-JUL-1989 (REL. 11, CREATED)
DT  01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  MAJOR CAPSID PROTEIN L1.
GN  L1.
OS  EUROPEAN ELK PAPILLOMAVIRUS (EPPV).
OC  VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPOVAVIRIDAE; PAPILLOMAVIRUSES.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  ERIKSSON A.;
RL  SUBMITTED (AUG-1987) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; M15953; G484027; -.
DR  PIR; G29499; P1WLEP
KW  COAT PROTEIN; LATE PROTEIN.
SQ  SEQUENCE 501 AA; 56459 MW; E4C2238D CRC32;

Query Match      69.2%; Score 45; DB 10; Length 501;
Best Local Similarity 75.0%; Pred. No. 8.79e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db  11 ylpptpvt 18
    || | |||
QY  1 YLEPGPVT 8

RESULT 11
ID  VLI_PAPVD      STANDARD;      PRT;    513 AA.
AC  P03104;
DT  21-JUL-1986 (REL. 01, CREATED)
DT  21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  MAJOR CAPSID PROTEIN L1.
GN  L1.
OS  DEER PAPILLOMAVIRUS.
OC  VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPOVAVIRIDAE; PAPILLOMAVIRUSES.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 85293253.
RA  GROFF D.E., LANCASTER W.D.;
RL  J. VIROL. 56:85-91(1985).
DR  EMBL; M11910; G808795; -.

DR  PIR; A03645; P1WLDP.
KW  COAT PROTEIN; LATE PROTEIN.
SQ  SEQUENCE 513 AA; 58145 MW; 0798B7F0 CRC32;

Query Match      69.2%; Score 45; DB 10; Length 513;
Best Local Similarity 75.0%; Pred. No. 8.79e+00;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db  11 ylpptpvt 18
    || | |||
QY  1 YLEPGPVT 8

RESULT 12
ID  NRML_CHICK     STANDARD;      PRT;    555 AA.
AC  P51027;
DT  01-OCT-1996 (REL. 34, CREATED)
DT  01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  NATURAL RESISTANCE-ASSOCIATED MACROPHAGE PROTEIN 1 (NRAMP 1).
GN  NRAMP1.
OS  GALLUS GALLUS (CHICKEN).
OC  EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC  GALLIFORMES.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-WHITE LEHORN; TISSUE-SPLEEN;
RX  MEDLINE; 96231835.
RA  HU J., BUMSTEAD N., SKAMENE E., GROS P., MALO D.;
RL  DNA CELL BIOL. 15:113-123(1996).
CC  -!- FUNCTION: MACROPHAGE-SPECIFIC MEMBRANE TRANSPORT FUNCTION.
CC  CONTROLS NATURAL RESISTANCE TO INFECTION WITH INTRACELLULAR
CC  PARASITES (BY SIMILARITY).
CC  -!- INDUCTION: IN RESPONSE TO LYMPHOKINE OR BACTERIAL PRODUCTS.
CC  -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC  -!- TISSUE SPECIFICITY: MACROPHAGES; SPLEEN AND THYMUS AND AT LOWER
CC  LEVEL IN LIVER AND LUNG.
CC  -!- SIMILARITY: BELONGS TO THE NRAMP FAMILY.
EMBL; U40598; G1222533; -.
KW  TRANSPORT; TRANSMEMBRANE; GLYCOPROTEIN; MACROPHAGE.
FT  DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 64 81 EXTRACELLULAR (POTENTIAL).
FT  DOMAIN 82 90 POTENTIAL.
FT  TRANSMEM 91 110 CYTOPLASMIC (POTENTIAL).
FT  DOMAIN 111 147 POTENTIAL.
FT  TRANSMEM 148 168 EXTRACELLULAR (POTENTIAL).
FT  DOMAIN 169 172 POTENTIAL.
FT  TRANSMEM 173 192 EXTRACELLULAR (POTENTIAL).
FT  DOMAIN 193 201 POTENTIAL.
FT  TRANSMEM 202 222 CYTOPLASMIC (POTENTIAL).
FT  DOMAIN 223 245 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 246 264 POTENTIAL.
FT  DOMAIN 265 292 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 293 312 POTENTIAL.
FT  DOMAIN 313 354 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 355 374 POTENTIAL.
FT  DOMAIN 375 405 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 406 423 POTENTIAL.
FT  DOMAIN 424 434 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 435 455 POTENTIAL.
FT  DOMAIN 456 471 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 472 493 POTENTIAL.
FT  DOMAIN 494 501 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 502 521 POTENTIAL.
FT  DOMAIN 522 555 CYTOPLASMIC (POTENTIAL).
FT  CARBOHYD 329 329 POTENTIAL.
FT  CARBOHYD 343 343 POTENTIAL.
SQ  SEQUENCE 555 AA; 60966 MW; C5249DFB CRC32;

Query Match      69.2%; Score 45; DB 6; Length 555;
Best Local Similarity 71.4%; Pred. No. 8.79e+00;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

Db 77 yldpgpv 83
  1 YLEPGPV 7
  1 YLEPGPV 7

RESULT 13
ID NDKB_HUMAN STANDARD; PRT; 152 AA.
AC P22392;
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE NUCLEOSIDE DIPHOSPHATE KINASE B (EC 2.7.4.6) (NDK B) (NDP KINASE B)
DE (NM23-H2) (C-MYC PURINE-BINDING TRANSCRIPTION FACTOR PUF).
GN NME2 OR NM23B.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
  [1]
RN SEQUENCE, SUBUNITS, AND ACTIVE SITE.
RX MEDLINE; 91224972.
RX GILLES A.-M., PRESPECAN E., VONICA A., LASCU I.;
RL J. BIOL. CHEM. 266:8784-8789(1991).
  [2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 91105674.
RA STAHL J.A., LEONE A., ROSENGARD A.M., PORTER L., LIOTTA L.A.,
RA STEEG P.S., KING C.R.;
RL CANCER RES. 51:445-449(1991).
  [3]
RN SEQUENCE FROM N.A.
RX MEDLINE; 93324921.
RA POSTEL E.H., BERBERICH S.J., FLINT S.J., FERRONE C.A.;
RL SCIENCE 261:478-480(1993).
  [4]
RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE; 95387396.
RA WEBB P.A., PERISIC O., MENDOLA C.E., BACKER J.M., WILLIAMS R.L.;
RL J. MOL. BIOL. 251:574-587(1995).
CC -1- FUNCTION: MAJOR ROLE IN THE SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATES
CC -1- OTHER THAN ATP.
CC -1- FUNCTION: ACTS AS A TRANSCRIPTIONAL ACTIVATOR OF THE C-MYC GENE;
CC BINDS DNA NONSPECIFICALLY (ACCORDING TO REF. 3).
CC -1- CATALYTIC ACTIVITY: ATP + NUCLEOSIDE = ADP + NUCLEOSIDE
CC TRIPHOSPHATE.
CC -1- SUBUNIT: HEXAMER OF TWO DIFFERENT CHAINS: A AND B (A6, A5B, ...,
CC A5, B6).
CC -1- DISEASE: THIS PROTEIN IS FOUND IN REDUCED AMOUNT IN TUMOUR CELLS
CC OF HIGH METASTATIC POTENTIAL.
CC -1- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC.
CC -1- SIMILARITY: TO OTHER NDK ISOZYMES.
DR EMBL; X58965; -, NOT_ANNOTATED_CDS.
DR EMBL; M36981; G189240; -.
DR EMBL; L16785; G349476; -.
DR PIR; B39838; B39838.
DR HSP; P08879; INDL.
DR MIM; 156491; -.
DR PROSITE; PS00469; NDP_KINASES.
KW TRANSFERASE; KINASE; ATP-BINDING; NUCLEAR PROTEIN; ANTI-ONCOGENE;
KW DNA-BINDING; TRANSCRIPTION REGULATION; ACTIVATOR.
FT MOD_RES 22 22 BLOCKED.
FT ACT_SITE 118 118
SQ SEQUENCE 152 AA; 17298 MW; 77B37CEA CRC32;

Query Match 67.7%; Score 44; DB 6; Length 152;
Best Local Similarity 55.6%; Pred. No. 1.46e+01;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 67 ymnsqpvva 75
  1 YLEPGPVTA 9
  1 YLEPGPVTA 9

RESULT 14
ID NDKB_RAT STANDARD; PRT; 152 AA.
AC P19804;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE NUCLEOSIDE DIPHOSPHATE KINASE B (EC 2.7.4.6) (NDK B) (NDP KINASE B)
DE (P18).
GN NME2.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
  [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN-WISTAR;
RX MEDLINE; 90368787.
RA KIMURA N., SHIMADA N., NOMURA K., WATANABE K.;
RL J. BIOL. CHEM. 265:15744-15749(1990).
  [2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 92332552.
RA ISHIKAWA N., SHIMADA N., MUNAKATA Y., WATANABE K., KIMURA N.;
RL J. BIOL. CHEM. 267:14366-14372(1992).
  [3]
RN SEQUENCE FROM N.A.
RX STRAIN-SPRAGUE-DRAWLEY; TISSUE-MAST CELLS;

```



RX MEDLINE: 92256389.  
RA HEMMERICH S., YARDEN Y., PECHT I.:  
RL BIOCHEMISTRY 31:4574-4579(1992).  
CC -!- FUNCTION: MAJOR ROLE IN THE SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATES  
CC OTHER THAN ATP.  
CC -!- CATALYTIC ACTIVITY: ATP + NUCLEOSIDE = ADP + NUCLEOSIDE  
CC TRIPHOSPHATE.  
CC -!- SUBUNIT: HEXAMER OF TWO DIFFERENT CHAINS: A AND B (A6, A5B, ...,  
CC A5, B6).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND PLASMA MEMBRANE.  
CC -!- SIMILARITY: TO OTHER NDK ISOZYMES.  
DR EMBL; M55331; G205662; -.  
DR EMBL; M91597; G206580; -.  
DR PIR; A38369; A38369.  
DR PIR; A41849; A41849.  
DR PIR; A42967; A42967.  
DR HSSP; P08879; 1NDL.  
DR PROSITE; PS00469; NDP\_KINASES.  
KW TRANSFERASE; KINASE; ATP-BINDING.  
FT MOD\_RES ?2 ?2 BLOCKED.  
FT ACT\_SITE 118 118  
FT CONFLICT 89 89 V -> W (IN REF. 3).  
SQ SEQUENCE 152 AA; 17283 MW; 2E2FC37F CRC32;

Query Match 67.7%; Score 44; DB 6; Length 152;  
Best Local Similarity 55.6%; Pred. No. 1.46e+01;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 67 ymngpvva 75  
|::|||  
Qy 1 YLEPGVTA 9

Search completed: Tue Jun 10 11:31:00 1997  
Job time : 10 secs.

\*\*\*\*\*

W P S R E H (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:33:02 1997; MasPar time 1.95 Seconds  
Tabular output not generated. 50.250 Million cell updates/sec

Title: >US-08-231-565A-39  
Description: (1-9) from US08231565A.pap  
Perfect Score: 52  
Sequence: 1 LNVSLADTN 9

Scoring table: PAM 150  
Gap 15

Searched: 92623 seqs, 10896596 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq26  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19

Statistics: Mean 14.488; Variance 42.223; scale 0.343

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	52	100.0	9	15	gp100 melanoma antige	8.45e+00
2	52	100.0	661	15	MART-1 melanoma antig	8.45e+00
3	52	100.0	661	15	MART-1 melanoma antig	8.45e+00
4	52	100.0	661	14	Melanoma associated a	8.45e+00
5	46	88.5	61	19	Human cytochrome P450	4.20e+01
6	44	84.6	3038	14	Aspergillus terreus t	7.09e+01
7	42	80.8	2227	3	Attenuated hepatitis	1.19e+02
8	42	80.8	2227	3	Sequence of viral L43	1.19e+02
9	40	76.9	1252	14	B. sphaericus SLP.	1.97e+02
10	40	76.9	1810	16	Chicken cytotactin.	1.97e+02
11	39	75.0	323	18	Carbonyl reductase.	2.54e+02
12	39	75.0	567	1	Gamma-glutamyl transp	2.54e+02
13	38	73.1	23	15	C.albicans enolase pe	3.25e+02
14	38	73.1	157	15	Recombinant C.albican	3.25e+02
15	38	73.1	440	15	Recombinant C.albican	3.25e+02
16	38	73.1	482	8	Lapine gran-negative	3.25e+02
17	38	73.1	482	10	Lapine polysaccharide	3.25e+02
18	38	73.1	1541	1	Igal protease.	3.25e+02
19	37	71.2	51	4	Unidentified protein	4.16e+02
20	37	71.2	273	4	Ospa polypeptide.	4.16e+02

21	37	71.2	374	19	W01125	Lycopene cyclase.	4.16e+02
22	37	71.2	374	3	R13987	Lycopene cyclase - va	4.16e+02
23	37	71.2	993	1	R05935	Secreted gp10b subuni	5.31e+02
24	36	69.2	110	11	R56390	Extracellular hGHR-CD	5.31e+02
25	36	69.2	112	11	R56391	Extracellular hGHR-CD	5.31e+02
26	36	69.2	144	3	R15322	IL-2 chimeric antibod	5.31e+02
27	36	69.2	144	3	R15321	IL-2 chimeric antibod	5.31e+02
28	36	69.2	148	16	R77017	Partial human ISS pat	5.31e+02
29	36	69.2	205	10	R54817	CHEP-3.	5.31e+02
30	36	69.2	237	4	R22229	Truncated human growt	5.31e+02
31	36	69.2	246	11	R56389	Human growth hormone	5.31e+02
32	36	69.2	249	2	R06867	Hormone binding regio	5.31e+02
33	36	69.2	269	1	R05045	Soluble human growth	5.31e+02
34	36	69.2	372	4	R22239	Human ZP3.	5.31e+02
35	36	69.2	387	16	R80227	Rabbit 50 kDa dystrop	5.31e+02
36	36	69.2	397	8	R43586	Ragweed Pollen Allerg	5.31e+02
37	36	69.2	637	1	P92108	Human growth hormone	5.31e+02
38	36	69.2	638	1	P81327	Rabbit growth hormone	5.31e+02
39	36	69.2	638	1	P81326	Human growth hormone	5.31e+02
40	36	69.2	638	1	P92107	Rabbit growth hormone	5.31e+02
41	36	69.2	712	4	R27823	Sequence encoded by T	5.31e+02
42	36	69.2	715	19	R97978	cadA gene product of	5.31e+02
43	36	69.2	717	4	R27824	Sequence encoded by T	5.31e+02
44	36	69.2	886	1	P80345	Sequence encoded by S	5.31e+02
45	36	69.2	2466	13	R71498	Human protein tyrosin	5.31e+02

## ALIGNMENTS

RESULT 1  
ID R84207 standard; Peptide; 9 AA.  
AC R84207;  
DE 25-APR-1996 (first entry)  
DE gp100 melanoma antigen immunogenic peptide (G10-400).  
KW gp100; melanoma antigen recognised by T-cells; MART; melanoma;  
KW metastatic melanoma; tumour-associated antigen;  
KW immunogenic peptide; diagnosis; prognosis; prophylaxis;  
KW therapy; vaccine.  
OS Synthetic.  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
DR WPI, 95-382963/49.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 55; Page 131; 184pp; English.  
CC The immunogenic peptide is derived from cDNA25 (R84854), a  
CC melanoma antigen derivative of gp100 (see R84855). The  
CC peptide and its derivatives (see R84200-R84211) are used in  
CC medicaments (vaccines) for the treatment or prevention (by  
CC immunization) of melanoma. Antibodies against melanoma-specific  
CC antigens and its immunogenic peptides may be used in the  
CC detection and isolation of the antigen from a sample, the  
CC detection of which is indicative of a disease state  
CC (melanoma or metastatic melanoma).  
SQ Sequence 9 AA;

Query Match 100.0%; Score 52; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.45e+00;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 Invsladtn 9

Oy 1 LNVSLADTN 9

RESULT 2  
ID R84855 standard; Protein; 661 AA.

R84855;  
AC 08-MAY-1996 (revised)  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen gp100.  
KW gp100; MART-1; melanoma antigen recognised by T-cell;  
KW cDNA25 antigen derivative; melanocyte; melanoma;  
KW metastatic melanoma; tumour-associated antigen; immunogen;  
KW diagnosis; prognosis; prophylaxis; therapy; vaccine.  
OS Mammalian sp.  
FH Key Location/Qualifiers  
FT Peptide 154..163  
FT /label= G9-154\_immunogenic\_peptide  
FT Peptide 208..217  
FT /label= G9-209\_immunogenic\_peptide  
FT /note= "see R84210"  
FT Peptide 280..288  
FT /label= G9-280\_immunogenic\_peptide  
FT /note= "see R84208"  
FT Peptide 457..266  
FT /label= immunogenic\_peptide  
FT Peptide 476..485  
FT /label= immunogenic\_peptide  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.  
PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
PI WPI: 95-382963/49.  
DR DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 81; Fig 7A; 184pp; English.  
CC gp100 is a melanoma-specific antigen (MART-1) which  
CC is recognized by T-lymphocytes, and is a derivative of the  
CC melanoma-specific antigen cDNA25 (see R84854). gp100 is a source  
CC of immunogenic peptides which are optionally modified to enhance  
CC their binding to a MHC molecule, and used in medicaments,  
CC especially vaccines, for the treatment or prevention (by  
CC immunisation) of melanoma. Antibodies against cDNA2 and its  
CC immunogenic peptides may be used in the detection and isolation  
CC of the antigen from a sample, the detection of which is indicative  
CC of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 661 AA;  
  
Query Match 100.0%; Score 52; DB 15; Length 661;  
Best Local Similarity 100.0%; Pred. No. 8.45e+00;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 567 lmvsladtn 575  
QY 1 LNVSLADTN 9  
|||||  
RESULT 3  
ID R84854 standard; Protein; 661 AA.  
AC R84854;  
DT 08-MAY-1996 (revised)  
DT 20-APR-1996 (first entry)  
DE MART-1 melanoma antigen cDNA25.  
KW cDNA25; MART-1; melanoma antigen recognised by T-cell;  
KW gp100 antigen derivative; melanoma; metastatic melanoma;  
KW tumour-associated antigen; immunogen; diagnosis; prognosis;  
KW prophylaxis; therapy; vaccine.  
OS Mammalian sp.  
FH Key Location/Qualifiers  
FT Peptide 457..466  
FT /label= antigenic\_peptide  
FT /note= "see R84199"  
PN W09529193-A2.  
PD 02-NOV-1995.  
PF 21-APR-1995; U05063.

PR 22-APR-1994; US-231565.  
PR 05-APR-1995; US-417174.  
PA (USSH ) US SEC DEPT HEALTH.  
PI Kawakami Y, Rosenberg SA;  
PI WPI: 95-382963/49.  
DR N-PSDB; T02716.  
PT DNA encoding melanoma antigens recognised by T-lymphocytes - also  
PT vectors, host cells and antibodies, used to detect, treat and  
PT immunise animal against melanoma.  
PS Claim 81; Fig 5A; 184pp; English.  
CC cDNA2 is a melanoma antigen (MART-1) which is recognized by  
CC T-lymphocytes, and is a derivative of the melanocyte-melanoma-  
CC specific antigen gp100 (see R84855). Antigen cDNA25 is a source  
CC of immunogenic peptides (see R84199) which are optionally modified  
CC (see R84200-R84211) to enhance their binding to a MHC molecule and  
CC used in medicaments, especially vaccines, for the treatment or  
CC prevention (by immunisation) of melanoma. Antibodies against cDNA2  
CC and its immunogenic peptides may be used in the detection and  
CC isolation of the antigen from a sample, the detection of which is  
CC indicative of a disease state (melanoma or metastatic melanoma).  
SQ Sequence 661 AA;  
  
Query Match 100.0%; Score 52; DB 15; Length 661;  
Best Local Similarity 100.0%; Pred. No. 8.45e+00;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 567 lmvsladtn 575  
QY 1 LNVSLADTN 9  
|||||  
RESULT 4  
ID R78646 standard; Protein; 661 AA.  
AC R78646;  
DT 22-JAN-1996 (first entry)  
DE Melanoma associated antigen gp100.  
KW Melanoma; antigen; vaccine; immunogen; primer; probe; detection;  
KW identification; tumour; gp100.  
OS Homo sapiens.  
PN EP-668350-A1.  
PD 23-AUG-1995.  
PF 14-FEB-1995; 200348.  
PR 16-FEB-1994; EP-200337.  
PR 21-DEC-1994; EP-203709.  
PA (ALKU ) AKZO NOBEL NV.  
PI Adema GJ, Figdor CG;  
PI WPI: 95-284790/38.  
DR N-PSDB; Q96055.  
PT Melanoma associated antigen gp100 - used in vaccines and for the  
PT detection of tumours  
PS Claim 1; Page 22-24; 40pp; English.  
CC Immunogenic peptides derived from the melanoma associated antigen  
CC may be used in the production of vaccines. Nucleotide sequences  
CC encoding the immunogenic peptides may be used as primers and probes  
CC in the detection of melanoma cells. Tumour infiltrating lymphocytes  
CC capable of binding to the melanoma associated antigen can be  
CC cultured ex vivo and returned to melanoma particles, and when  
CC radiolabelled, they may be used to identify tumour deposits.  
SQ Sequence 661 AA;  
  
Query Match 100.0%; Score 52; DB 14; Length 661;  
Best Local Similarity 100.0%; Pred. No. 8.45e+00;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 567 lmvsladtn 575  
QY 1 LNVSLADTN 9  
|||||  
RESULT 5  
ID W05026 standard; Protein; 61 AA.  
AC W05026;  
DT 04-DEC-1996 (first entry)

DE Human cytochrome P450 IIC19 exon5/exon7 aberrant splice junction.  
 KW Aberrant splicing variant; abnormal transcription; exon 6; exon 7;  
 OS PCR; polymerase chain reaction; amplification; detection.  
 FH Homo sapiens.  
 FT Key Location/Qualifiers  
 FT Region 1..12  
 FT /note= "encoded by 3'-region of exon 5"  
 FT Region 13..61  
 FT /note= "encoded by 5'-region of exon 7"  
 FT J08168400-A.  
 PD 02-JUL-1996.  
 PE 19-DEC-1994; 315199.  
 PR 19-DEC-1994; JP-315199.  
 PA (SUMO ) SUMITOMO CHEM CO LTD.  
 DR WPI; 96-357254/36.  
 DR N-PSDB; T41565.  
 PT Detecting mutation(s) in human-originated cytochrome P450 IIC19 gene  
 PT - by amplifying exon 6 and/or 7 by PCR  
 FS Disclosure; Fig 2; lipp; Japanese.  
 CC Oligonucleotide primers of 8-50 nucleotides in length which are used  
 CC to amplify exon 6 and/or exon 7 of the human cytochrome P450 IIC19  
 CC gene are claimed. In some individuals, aberrant splicing of the gene  
 CC transcript can result in loss of exon 6 or loss of both exons 6 and 7.  
 CC The presence of splice variants can be detected by comparing the  
 CC length of PCR products generated from a particular template. The  
 CC present sequence is that encoded by the region around the splice  
 CC junction created by aberrant splicing of exon 5 to exon 7.  
 SQ Sequence 61 AA;  
 Query Match 88.5%; Score 46; DB 19; Length 61;  
 Best Local Similarity 87.5%; Pred. NO. 4.20e+01;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 19 lmvslaet 26  
 Qy 1 lmvsladt 8  
 RESULT 6  
 ID R74171 standard; Protein; 3038 AA.  
 AC R74171;  
 DT 18-JAN-1996 (first entry)  
 DE Aspergillus terreus triol polyketide synthase.  
 KW Triol polyketide synthase; TPKS; HMG-CoA reductase inhibitor;  
 KW hypercholesterolaemia; LDL- cholesterol.  
 OS Aspergillus terreus.  
 FH Key Location/Qualifiers  
 FT Region 181  
 FT /label= keto-acyl synthase motif  
 FT Region 654..658  
 FT /label= acetyl/malonyl transferase motif  
 FT Region 985..994  
 FT /label= dehydratase motif  
 FT Region 1446..1450  
 FT /label= methyl transferase motif  
 FT Region 1932..1937  
 FT /label= enoyl reductase motif  
 FT Region 2164..2169  
 FT /label= keto reductase motif  
 FT Region 2498  
 FT /label= acyl carrier protein motif  
 FT Misc\_difference 282..288  
 FT /label= misc feature  
 FT Misc\_difference 1450..1460  
 FT /label= misc feature  
 FT Misc\_difference 1603..1612  
 FT /label= misc feature  
 FT Misc\_difference 2521..2535  
 FT /label= misc feature  
 PN W09512661-A.  
 PD 11-MAY-1995.  
 PF 28-OCT-1994; U12423.  
 PR 02-NOV-1993; US-148132.

PA (MERI ) MERCK & CO INC.  
 PI Conder MJ, Davis CR, Hendrickson LE, Mcada PC, Rambosek J;  
 PI Reeves CD, Vinci VA;  
 OS WPI; 95-193816/25.  
 DR N-PSDB; Q92323.  
 PT Novel DNA encoding triol poly-ketide synthase - used to isolate and  
 PT identify homologues of triol poly-ketide synthase, and in the treatment  
 PT of hyper-cholesterolaemia  
 PS Claim 12; Figure 2; 107pp; English.  
 CC The full-length TPKS-encoding DNA in plasmid pLOA was  
 CC designated pTPKS100. Splicing of the introns from the DNA  
 CC sequence and translation of the 914 nt ORF results in a  
 CC protein of 3038 AAs (R74171) with a mol. wt. of 269,090  
 CC daltons. Inspection of the TPKS AA sequence for active  
 CC site residues and motifs known to be associated with  
 CC polyketide synthases and fatty acid synthase (FAS) activities  
 CC resulted in the identification of candidates for expected  
 CC sites (see FT). Except for the presence of a methyl transferase,  
 CC not present in FAS, the succession of activities on the  
 CC TPKS protein is the same as that observed for the rat FAS  
 CC protein.  
 SQ Sequence 3038 AA;  
 Query Match 84.6%; Score 44; DB 14; Length 3038;  
 Best Local Similarity 77.8%; Pred. No. 7.09e+01;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Db 2869 ltvgladt 2877  
 Qy 1 lmvsladt 9  
 RESULT 7  
 ID R05697 standard; protein; 2227 AA.  
 AC R05697;  
 DT 15-AUG-1990 (first entry)  
 DE Attenuated hepatitis A virus  
 KW Hepatitis A virus; vaccine; attenuated.  
 OS Hepatitis A virus; strain HM-175.  
 FH Key Location/Qualifiers  
 FT Region 1..23  
 FT /label=VP4 = 1A  
 FT Region 24..245  
 FT /label=VP2 = 1B  
 FT Region 246..491  
 FT /label=VP3 = 1C  
 FT Region 492..791  
 FT /label=VP1 = 1D  
 FT Region 792..980  
 FT /label=2A  
 FT Region 981..1087  
 FT /label=2B  
 FT Region 1088..1422  
 FT /label=2C  
 FT Region 1423..1496  
 FT /label=3A  
 FT Region 1497..1519  
 FT /label=3B = VPg  
 FT Region 1520..1738  
 FT /label=3C  
 FT Region 1739..2227  
 FT /label=3D  
 PN US4894228-A.  
 PD 16-JAN-1990.  
 PE 12-JUL-1988; 217824.  
 PR 12-JUL-1988; US-217824.  
 PR 12-JUL-1988; US-652967.  
 PA (USSH) US Dept Health & Human.  
 PI Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstein SM,  
 PI Daemer RJ, Gust ID.  
 DR WPI; 90-075557/10.  
 DR N-PSDB; Q03512.  
 PT Vaccine against hepatitis A virus infection - comprises novel

PT attenuated hepatitis A virus strain.  
 PS Claim 1; Fig 1; 18pp; English.  
 CC The attenuated HAV is useful for inducing protective immunity against  
 CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by  
 CC several nucleotide changes distributed throughout the genome, is  
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is  
 CC suitable for use as an HAV vaccine. It is noted that not all the changes  
 CC are necessary for attenuation and use as a vaccine.  
 SQ Sequence 2227 AA;

Query Match 80.8%; Score 42; DB 1; Length 2227;  
 Best Local Similarity 77.8%; Pred. No. 1.19e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lvnlnlakt 1375  
 |||:||||  
 QY 1 LNVSLADTN 9

RESULT 8  
 ID P60066 standard; Protein; 2227 AA.  
 AC P60066;  
 DT 26-JUN-1991 (first entry)  
 DE Sequence of viral I434 polypeptide encoded by the complete  
 DE nucleotide sequence of the HAV genome.  
 KW Diagnosis; vaccine; passive immunotherapy.  
 OS Hepatitis A virus.

FH Key Location/Qualifiers  
 FT Region 1..245

FT /label= P1.1A 246..491  
 FT Region 1B  
 FT /label= 1B  
 FT Region 492..836  
 FT /label= 1C  
 FT Region 837..980  
 FT /label= P2.2A 981..1076  
 FT Region 2B  
 FT /label= 2B 1077..1422  
 FT /label= 2C 1423..1484  
 FT Region P3.3A 1485..1507  
 FT Region 3B 1508..1678  
 FT /label= 3C 1679..2227  
 FT /label= 3D

FT EP-199480-A.  
 PD 29-OCT-1986.  
 PF 03-APR-1986; 302465.

PR 03-APR-1985; US-719329.  
 PA (CHIR-) CHIRON CORP.

PI Dina D, Potter SJ, Vannest GA, Caput D;  
 DR WPI: 86-286213/44.

DR N-PSDB; N60080.

PT Hepatitis A virus nucleotide sequence and polypeptide - and use  
 PT in prodn. of vaccines and diagnostic probes

PS Claim 5; Fig 1; 18pp; English.  
 CC N60080 and oligonucleotide fragments are useful in detection of  
 CC hepatitis A virus; transmits useful in vaccines without risk of  
 CC polypeptides and fragments useful in vaccines which are capable  
 CC infection by the virus or in prodn. of particles which are capable  
 CC of inducing immunocompetent B cells for passive immunotherapy. Pref.  
 CC epitope is derived from AAs 445-657 or 792-848 of the HAV  
 CC polypeptide sequence (P60066).  
 SQ Sequence 2227 AA;

Query Match 80.8%; Score 42; DB 3; Length 2227;  
 Best Local Similarity 77.8%; Pred. No. 1.19e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lvnlnlakt 1375

|||:||||  
 QY 1 LNVSLADTN 9

RESULT 9  
 ID R80530 standard; Protein; 1252 AA.

AC R80530;  
 DT 22-DEC-1995 (first entry)

DE B. sphaericus SLP.  
 KW Surface layer protein; SLP; fusion protein; vaccine; antigen;  
 KW surface expression; epitope.

OS Bacillus sphaericus.

FH Key Location/Qualifiers

FT Peptide 1..30

FT /label= sig\_peptide

PN W09519371-AZ.

PD 20-JUL-1995.

PF 13-JAN-1995; E00147.

PR 14-JAN-1994; GB-000650.

PA (SOLV ) SOLVAY SA.

PI Deblaere RY, Desomer J, Dhaese P;

DR WPI: 95-263827/34.

DR N-PSDB; Q99430.

PT Host cell expressing surface layer protein fusion protein - used for  
 PT host presentation of antigens and vaccine prodn.

PS Disclosure; Fig.6; 95pp; English.

CC A probe based on the N-terminal sequence of B. sphaericus P-1 (LMG

CC P-13855) surface layer protein was used to screen an HindIII-

CC generated library to isolate the slp gene. Promoter regions

CC of the gene are used in genetic constructs providing surface

CC expression of heterologous proteins in P-1 hosts.

SQ Sequence 1252 AA;

Query Match 76.9%; Score 40; DB 14; Length 1252;  
 Best Local Similarity 55.6%; Pred. No. 1.97e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 943 lgislada 951

|||:||||  
 QY 1 LNVSLADTN 9

RESULT 10

ID R94563 standard; Protein; 1810 AA.

AC R94563;

DT 21-JUN-1996 (first entry)

DE Chicken cytotoxin.

KW Cytotoxin; neuron; neurite; cell attachment; cell elongation;

KW fibronectin.

OS Gallus sp.

FH Key Location/Qualifiers

FT Region 1..591

FT /note= "EGF-like repeat region"

FT Region 592..773

FT /note= "fibronectin type III repeats I-II"

FT Region 774..864

FT /note= "fibronectin type III repeat III"

FT Region 866..956

FT /note= "fibronectin type III repeat IV"

FT Region 957..1044

FT /note= "fibronectin type III repeat V"

FT Region 1318..1398

FT /note= "fibronectin type III repeat VI"

FT Region 1045..1317

FT /note= "alternatively spliced fibronectin type

FT III repeat VaVbVc"

FT Region 1132..1569

FT /note= "fibronectin type III repeats VII and

FT VIII (lacking 13 N-terminal amino acids)"

FT Region 1318..1398

FT /note= "fibronectin type III repeats III-VI"

FT Region 1570..1810

FT /note= "fibrinogen domain plus fibronectin type

```

FT III repeat VIII N-terminus"
PN W09608513-A1.
PD 21-MAR-1996.
PF 14-SEP-1995; U11684.
PR 16-SEP-1994; US-308359.
PA (SCRI ) SCRIPPS RES INST.
PI Crossin KL, Phillips G, Prieto AL;
DR WPI; 96-179904/18.
DR N-PSDB; T14548.
DR Cytotactin polypeptide(s), derivs. and antibodies - capable of
PT stimulating neuronal cell attachment, neurite out-growth and cell
PT elongation
PS Claim 9; Page 118-126; 159pp; English.
CC Chicken cytotactin (R94563) and human cytotactin (R94562) are
CC multidomain extracellular matrix proteins capable of stimulating
CC neuronal cell attachment, neurite outgrowth and cell elongation.
CC The cytotactins, or functional portions of them (see R94564-69),
CC can be obtd. by expression of encoding DNA (see T14548 and
CC (T14547) in host cells, e.g. as GST fusion proteins in Escherichia
CC coli. They have therapeutic and diagnostic applns. and can be
CC used to raise antibodies.
SQ Sequence 1810 AA;

Query Match 76.9%; Score 40; DB 16; Length 1810;
Best Local Similarity 55.6%; Pred. No. 1.97e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 37 Invltlpedn 45
Qy 1 LNVSLADTN 9

RESULT 11
ID R96294 standard; Protein; 323 AA.
AC R96294;
DT 18-NOV-1996 (first entry)
DE Carbonyl reductase.
KW carbonyl reductase; R-gamma-substituted-beta-hydroxybutyrate;
KW gamma-substituted acetacetate; raw material; drug synthesis;
KW agricultural chemical production.
OS Sporobolomyces salmonicolor IF01038.
FH Key Location/Qualifiers
FT /label= mature_protein
FT /label= 2..323
PN J08103269-A.
PD 23-APR-1996.
PF 07-OCT-1994; 244090.
PR 07-OCT-1994; JP-244090.
PA (ELED ) DENKI KAGAKU KOGYO KK.
DR WPI; 96-253873/26.
DR N-PSDB; T29139.
PT Carbonyl reductase gene and protein - used in the production of
PT R-gamma-substituted-beta-hydroxybutyrate, for use in drugs and
PT agrochemicals
PS Claim 1; Page 9-10; 14pp; Japanese.
CC The present sequence is that of a carbonyl reductase (ALD) isolated from
CC Sporobolomyces salmonicolor IF01038. The enzyme is used in the synthesis
CC of an (R)-gamma-substituted-beta-hydroxybutyrate (GSBH) from a gamma-
CC substituted acetacetate. The GSBH is useful as a raw material for the
CC synthesis of drugs and agricultural chemicals.
SQ Sequence 323 AA;

Query Match 75.0%; Score 39; DB 18; Length 323;
Best Local Similarity 75.0%; Pred. No. 2.54e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 135 levsldvt 142
Qy 1 LNVSLADT 8

RESULT 12
ID R04229 standard; protein; 567 AA.

```

```

AC R04229;
DT 17-SEP-1990 (first entry)
DE Gamma-glutamyl transpeptidase gene.
KW Gamma-glutamyl transpeptidase; GTP; cancer; ds.
OS Homo sapiens.
PN J02079971-A.
PD 20-MAR-1990.
PF 16-SEP-1988; 229772.
PR 16-SEP-1988; JP-229772.
PA (ORIY) Oriental Yeast KK.
DR WPI; 90-134490/18.
DR N-PSDB; Q03692.
PT Primary structure of human gamma-glutamyl transpeptidase -
PT having strong activity against various carcinoma.
PS Claim 5; Fig 3; 9pp; Japanese.
SQ Sequence 567 AA;

Query Match 75.0%; Score 39; DB 1; Length 567;
Best Local Similarity 62.5%; Pred. No. 2.54e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 263 lnislqda 270
Qy 1 LNVSLADT 8

RESULT 13
ID R81710 standard; peptide; 23 AA.
AC R81710;
DT 10-MAY-1996 (first entry)
DE C.albicans enolase peptide #1.
KW Candida albicans; enolase; lambda-gt11; serum; antibody; candidiasis;
KW thrush.
OS Synthetic.
PN W09526362-A1.
PD 05-OCT-1995.
PF 27-MAR-1995; AU0176.
PR 25-MAR-1994; AU-004732.
PA (UYCU-) UNIV CURTIN TECHNOLOGY.
PI Franklyn KW, Warrington JR;
DR WPI; 95-351294/45.
PT Protein or peptide reactive with Candida albicans induced antibodies
PT - also DNA and vectors, used to treat, prevent or diagnose C.
PT albicans infections
PS Claim 7; Page 26; 37pp; English.
CC Peptides R81710-25 are derived from the Candida albicans enolase protein
CC (R81708). The peptides can be used to diagnose, in a serum sample,
CC antibodies induced by a Candida albicans infection and hence the
CC infection itself. The gene encoding the enolase was isolated from a
CC C.albicans cDNA expression library in lambda-gt11 using human sera from a
CC patient with a clinical history of vaginosis. The resultant clone
CC contained a 470 bp open reading frame encoding residues 93-249 (R81709).
CC This clone was used to screen a genomic library and obtain the full
CC length clone.
SQ Sequence 23 AA;

Query Match 73.1%; Score 38; DB 15; Length 23;
Best Local Similarity 66.7%; Pred. No. 3.25e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 15 lgvslaaan 23
Qy 1 LNVSLADTN 9

RESULT 14
ID R81709 standard; peptide; 157 AA.
AC R81709;
DT 10-MAY-1996 (first entry)
DE Recombinant C.albicans enolase residues 93-249.
KW Candida albicans; enolase; lambda-gt11; serum; antibody; candidiasis;
KW thrush.
OS Candida albicans.

```

Job time : 9 secs.

PN WO9526362-AL.  
 PD 05-OCT-1995.  
 PF 27-MAR-1995; AU0176.  
 PR 25-MAR-1994; AU-004732.  
 PA (YUCU-) UNIV CURTIN TECHNOLOGY.  
 PI Franklyn KM, Warrington JR;  
 DR WPI: 95-351294/45.  
 PT Protein or peptide reactive with Candida albicans induced antibodies  
 PT - also DNA and vectors, used to treat, prevent or diagnose C.  
 PT albicans infections  
 PS Claim 6; Page 23-24; 37pp; English.  
 CC The amino acid sequence of residues 93-249 of the 48 kD Candida albicans  
 CC enolase. This DNA encoding this fragment of the full length enolase was  
 CC isolated from a C.albicans cDNA expression library in lambda-gt11 using  
 CC human sera from a patient with a clinic history of vaginosis. The  
 CC resultant clone contained a 470 bp open reading frame encoding residues  
 CC 93-249. This clone was used to screen a genomic library and obtain the  
 CC full length clone. The protein sequence was used to generate the  
 CC peptides R81710-25 which can be used to diagnose, in a serum sample,  
 CC antibodies induced by a candida albicans infection and hence the  
 CC infection itself.  
 SQ Sequence 157 AA;

Query Match 73.1%; Score 38; DB 15; Length 157;  
 Best Local Similarity 66.7%; Pred. NO. 3.25e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 21 lgvslaaan 29  
 | ||||| :  
 QY 1 LNVSLADTN 9

RESULT 15  
 ID R81708 standard; Protein; 440 AA.  
 AC R81708;  
 DT 10-MAY-1996 (first entry)  
 DE Recombinant C.albicans enolase.  
 KW Candida albicans; enolase; lambda-gt11; serum; antibody; candidiasis;  
 OS Candida albicans.  
 PN WO9526362-AL.  
 PD 05-OCT-1995.  
 PF 27-MAR-1995; AU0176.  
 PR 25-MAR-1994; AU-004732.  
 PA (YUCU-) UNIV CURTIN TECHNOLOGY.  
 PI Franklyn KM, Warrington JR;  
 DR WPI: 95-351294/45.  
 DR N-PSDB; T00755.  
 PT Protein or peptide reactive with Candida albicans induced antibodies  
 PT - also DNA and vectors, used to treat, prevent or diagnose C.  
 PT albicans infections  
 PS Claim 4; Page 23-24; 37pp; English.  
 CC The amino acid sequence of the candida albicans enolase which has a mol.  
 CC wt. of 48 kD. The gene encoding was isolated from a C.albicans cDNA  
 CC expression library in lambda-gt11 using human sera from a patient with a  
 CC clinical history of vaginosis. The resultant clone contained a 470 bp  
 CC open reading frame encoding residues 93-249 (R81709). This clone was  
 CC used to screen a genomic library and obtain the full length clone. The  
 CC protein sequence was used to generate the peptides R81710-25 which can be  
 CC used to diagnose, in a serum sample, antibodies induced by a Candida  
 CC albicans infection and hence the infection itself.  
 SQ Sequence 440 AA;

Query Match 73.1%; Score 38; DB 15; Length 440;  
 Best Local Similarity 66.7%; Pred. NO. 3.25e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 113 lgvslaaan 121  
 | ||||| :  
 QY 1 LNVSLADTN 9

Search completed: Tue Jun 10 11:33:11 1997





```

ORGANISM      #formal_name Homo sapiens #common_name man
DATE          19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change
ACCESSIONS    30-Sep-1993
REFERENCE      A41234
#authors      Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.;
              Gilbert, D.J.; Jenkins, N.; Barton, D.; Francke, U.;
              Kobayashi, Y.; Kim, K.K.
#journal      Proc. Natl. Acad. Sci. U.S.A. (1991) 88:9228-9232
#title        A melanocyte-specific gene, Pmel 17, maps near the silver
              coat color locus on mouse chromosome 10 and is in a
              syntenic region on human chromosome 12.
#cross-references MIM:92021023
#accession    A41234
#status       preliminary
#molecule_type mRNA
#residues     1-668 #label KWO
#cross-references GB:M77348
SUMMARY       #length 668 #molecular-weight 70932 #checksum 6409
Query Match   100.0%; Score 52; DB 13; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.30e-01;
Matches       9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 567 lnvsldatn 575
|||||
QY 1 LNVS�ADTN 9

RESULT 3
ENTRY      A49179 #type fragment
TITLE      melanoma antigen homolog rpe1 - bovine (fragment)
ORGANISM    #formal_name Bos primigenius taurus #common_name cattle
DATE        19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
23-Mar-1995
ACCESSIONS  A49179
REFERENCE    A49179
#authors     Kim, R.Y.; Wistow, G.J.
#journal     Exp. Eye Res. (1992) 55:657-662
#title       The cDNA RPE1 and monoclonal antibody HMB-50 define gene
              products preferentially expressed in retinal pigment
              epithelium.
#cross-references MIM:93122163
#accession   A49179
#status       preliminary
#molecule_type nucleic acid
#residues    1-491 #label KIM
#cross-references NCBI:N122438; NCBI:P122439
#experimental_source retinal pigment epithelium
#note        sequence extracted from NCBI backbone
SUMMARY      #length 491 #checksum 3125
Query Match   94.2%; Score 49; DB 14; Length 491;
Best Local Similarity 88.9%; Pred. No. 1.10e+00;
Matches       8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 395 lnvsldatn 403
|||||
QY 1 LNVS�ADTN 9

RESULT 4
ENTRY      S53871 #type complete
TITLE      Pmel 17 protein - mouse
ORGANISM    #formal_name Mus musculus #common_name house mouse
DATE        27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change
03-Nov-1995
ACCESSIONS  S53871
REFERENCE    S53871
#authors     Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.;
              Chintamaneni, C.; Bennett, D.; Pickard, R.T.
#journal     Nucleic Acids Res. (1995) 23:154-158
#title       Mouse silver mutation is caused by a single base insertion in

```

---

```

#accession    S53871 the putative cytoplasmic domain of Pmel 17.
#status       preliminary
#molecule_type mRNA
#residues     #length 626 #molecular-weight 65979 #checksum 5710
SUMMARY       #length 626 #molecular-weight 65979 #checksum 5710
Query Match   94.2%; Score 49; DB 14; Length 626;
Best Local Similarity 88.9%; Pred. No. 1.10e+00;
Matches       8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 534 lnvsldatn 542
|||||
QY 1 LNVS�ADTN 9

RESULT 5
ENTRY      S04990 #type complete
TITLE      integrase - phage HK022
ORGANISM    #formal_name phage HK022
DATE        30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
18-Jun-1993
ACCESSIONS  S04990
REFERENCE    S04990
#authors     Yagil, E.; Dolev, S.; Oberto, J.; Kislev, N.; Ramaliah, N.;
              Weisberg, R.A.
#journal     J. Mol. Biol. (1989) 207:695-717
#title       Determinants of site-specific recombination in the lambdaoid
              coliphage HK022. An evolutionary change in specificity.
#cross-references MIM:89342457
#accession   S04990
#molecule_type DNA
#residues     1-357 #label YAG
#cross-references EMBL:X51962
GENETICS
#gene        int
#classification superclass phage lambda integrase
SUMMARY      #length 357 #molecular-weight 40358 #checksum 2665
Query Match   86.5%; Score 45; DB 8; Length 357;
Best Local Similarity 87.5%; Pred. No. 8.10e+00;
Matches       7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 251 lnsladt 258
|||||
QY 1 LNVS�ADT 8

RESULT 6
ENTRY      D30020 #type complete
TITLE      NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 - fruit
              fly (Drosophila yakuba) mitochondrion (SGC4)
ALTERNATE_NAMES NADH-ubiquinone oxidoreductase chain 1
ORGANISM        #formal_name mitochondrion Drosophila yakuba
DATE            05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change
11-Aug-1995
ACCESSIONS  D30020; M25797
REFERENCE    A92962
#authors     Clary, D.O.; Wolstenholme, D.R.
#journal     J. Mol. Evol. (1985) 22:252-271
#title       The mitochondrial DNA molecule of Drosophila yakuba:
              nucleotide sequence, gene organization, and genetic code.
#cross-references MIM:86089137
#accession   D30020
#molecule_type DNA
#residues     1-324 #label CLA
GENETICS
#genome       mitochondrion
#genetic_code SGC4
#start_codon  ATA
#classification superclass family NADH dehydrogenase (ubiquinone) chain 1
              membrane-associated complex; mitochondrion; NAD:
              oxidoreductase

```

```

SUMMARY      #length 324 #molecular-weight 37529 #checksum 422
Query Match  84.6%; Score 44; DB 5; Length 324;
Best Local Similarity 66.7%; Pred. No. 1.31e+01;
Matches      6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 190 ltlislaetn 198
| : |||||
| : |||||
QY 1 LNVSLADTN 9

RESULT 7
ENTRY S39630 #type complete
TITLE hypothetical protein B - Pseudomonas aeruginosa
ORGANISM #formal_name Pseudomonas aeruginosa
DATE 06-Mar-1994 #sequence_revision 06-Mar-1994 #text_change
06-Mar-1994
ACCESSIONS S39630
REFERENCE S39629
#authors Poole, K.; Heinrichs, D.E.; Neshat, S.
#journal Mol. Microbiol. (1993) 10:529-544
#title Cloning and sequence analysis of an EnvCD homologue in
Pseudomonas aeruginosa: regulation by iron and possible
involvement in the secretion of the siderophore pyoverdine.
#accession S39630
#status preliminary
#residues 1-1046 #label POO
SUMMARY #length 1046 #molecular-weight 112778 #checksum 6219

Query Match 80.8%; Score 42; DB 16; Length 1046;
Best Local Similarity 77.8%; Pred. No. 3.36e+01;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 738 lqvsladin 746
| : |||||
| : |||||
QY 1 LNVSLADTN 9

RESULT 8
ENTRY GNNYMK #type complete
TITLE genome polyprotein - human hepatitis A virus (strain
HM-175/7MK-5, attenuated HAV)
CONTAINS coat protein 1A; coat protein 1B; coat protein 1C; core
protein 2A; core protein 2B; core protein 2C; cysteine
proteinase (EC 3.4.22.-), protein 3C; protein 3A; protein
3B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
ORGANISM #formal_name human hepatitis A virus
#note host Homo sapiens (man)
DATE 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change
10-May-1996
ACCESSIONS A94149; A25914; A94508
REFERENCE A94149
#authors Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.;
Feinstone, S.M.; Purcell, R.H.
#journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:2497-2501
#title Complete nucleotide sequence of an attenuated hepatitis A
virus: comparison with wild-type virus.
#cross-references EMBL:M16632
#accession A94149
#status nucleic acid sequence not shown
#molecule_type genomic RNA
#residues 1-2227 #label COH
#note submitted to GenBank, August 1987
CLASSIFICATION #superfamily hepatitis A virus genome polyprotein
#superfamily hepatitis A virus genome polyprotein
KEYWORDS coat protein; core protein; cysteine proteinase; hydrolase;
nucleotidyltransferase; polyprotein
FEATURE
1-245 #product coat protein 1A #status predicted #label P1A\
246-491 #product coat protein 1B #status predicted #label P1B\
492-836 #product coat protein 1C #status predicted #label P1C\
837-980 #product core protein 2A #status predicted #label P2A\
981-1076 #product core protein 2B #status predicted #label P2B\

1077-1422 #product core protein 2C #status predicted #label P2C\
1423-1484 #product protein 3A #status predicted #label P3A\
1485-1507 #product cysteine proteinase, protein 3C #status
predicted #label P3C\
1508-1678 #product RNA-directed RNA polymerase, protein 3D #status
predicted #label P3D\
SUMMARY #length 2227 #molecular-weight 251515 #checksum 3810

Query Match 80.8%; Score 42; DB 4; Length 2227;
Best Local Similarity 77.8%; Pred. No. 3.36e+01;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lvnlnaktn 1375
| : |||||
| : |||||
QY 1 LNVSLADTN 9

RESULT 9
ENTRY GNNYHM #type complete
TITLE genome polyprotein - human hepatitis A virus (strain HM-175,
wild type)
CONTAINS coat protein 1A; coat protein 1B; coat protein 1C; coat
protein 1D; core protein 2A; core protein 2B; core protein
2C; cysteine proteinase (EC 3.4.22.-), protein 3C; protein
3A; protein 3B; RNA-directed RNA polymerase (EC 2.7.7.48),
protein 3D
ORGANISM #formal_name human hepatitis A virus
#note host Homo sapiens (man)
DATE 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change
10-May-1996
ACCESSIONS A25981
REFERENCE A25981
#authors Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White,
A.; Baroudy, B.M.
#journal J. Virol. (1987) 61:50-59
#title Complete nucleotide sequence of wild-type hepatitis A virus:
comparison with different strains of hepatitis A virus and
other picornaviruses.
#cross-references MUID:87061253
#accession A25981
#molecule_type genomic RNA
#residues 1-2227 #label COH
#cross-references EMBL:M14707
CLASSIFICATION #superfamily hepatitis A virus genome polyprotein
KEYWORDS coat protein; core protein; cysteine proteinase; hydrolase;
nucleotidyltransferase; polyprotein
FEATURE
1-23 #product coat protein 1A #status predicted #label VP4\
24-245 #product coat protein 1B #status predicted #label VP2\
246-491 #product coat protein 1C #status predicted #label VP3\
492-791 #product coat protein 1D #status predicted #label VP1\
792-980 #product core protein 2A #status predicted #label C2A\
981-1087 #product core protein 2B #status predicted #label C2B\
1088-1422 #product core protein 2C #status predicted #label C2C\
1423-1496 #product core protein 3A #status predicted #label C3A\
1497-1519 #product protein 3B #status predicted #label C3B\
1520-1738 #product cysteine proteinase, protein 3C #status
predicted #label C3C\
1739-2227 #product RNA-directed RNA polymerase, protein 3D #status
predicted #label C3D\
SUMMARY #length 2227 #molecular-weight 251506 #checksum 1963

Query Match 80.8%; Score 42; DB 4; Length 2227;
Best Local Similarity 77.8%; Pred. No. 3.36e+01;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lvnlnaktn 1375
| : |||||
| : |||||
QY 1 LNVSLADTN 9

RESULT 10

```



```
#accession A45502
#molecule_type DNA
#residues 1-178 ##label ASK
#note the authors translated the codon CAC for residue 173 as Gln
REFERENCE A31198
#authors Ingvarsson, S.; Asker, C.; Axelsson, H.; Klein, G.; Suemegi, J.
#journal Mol. Cell. Biol. (1988) 8:3168-3174
#title Structure and expression of B-myc, a new member of the myc gene family.
#cross-references MUID:89096904
#accession A31198
#molecule_type mRNA
#residues 1-120 ##label ING
COMMENT The sequence is related to the amino end of the myc protein. It is missing the region thought to be important for nuclear localization, DNA binding, transforming activity, and oligomerization.
GENETICS
#gene B-myc
CLASSIFICATION #superfamily transforming protein B-myc
KEYWORDS transforming protein
SUMMARY #length 178 #molecular-weight 19463 #checksum 5965
Query Match 76.9%; Score 40; DB 2; Length 178;
Best Local Similarity 66.7%; Pred. No. 8.31e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 3 lmvslangn 11
|:||||:|
QY 1 LNVSLADTN 9
RESULT 14
ENTRY JU0473 #type complete
TITLE hypothetical protein X19 (lysA region) - Bacillus subtilis
ORGANISM #formal_name Bacillus subtilis
DATE 23-Nov-1991 #sequence_revision 10-Feb-1995 #text_change 10-Feb-1995
ACCESSIONS S45536; JU0473
REFERENCE S45533
#authors Sorokin, A.; Zumstein, E.; Azevedo, V.; Ehrlich, S.D.; Serror, P.
#submission submitted to the EMBL Data Library, November 1993
#accession S45536
#molecule_type DNA
#residues 1-290 ##label SOR
#cross-references EMBL:L09228
REFERENCE JU0471
#authors Yamamoto, J.; Shimizu, M.; Yamane, K.
#journal Agric. Biol. Chem. (1991) 55:1615-1626
#title Molecular cloning and analysis of nucleotide sequence of the Bacillus subtilis lysA gene region using B. subtilis phage vectors and a multi-copy plasmid, pUB110.
#accession JU0473
#molecule_type DNA
#residues 113-290 ##label YAM
GENETICS
#start_codon TTG
SUMMARY #length 290 #molecular-weight 31294 #checksum 5892
Query Match 76.9%; Score 40; DB 10; Length 290;
Best Local Similarity 62.5%; Pred. No. 8.31e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 218 lmvltids 225
|:||||:|
QY 1 LNVSLADT 8
RESULT 15
ENTRY A64131 #type complete
```

```
TITLE arginine--trna ligase (EC 6.1.1.19) - Haemophilus influenzae (strain Rd KW20)
ALTERNATE_NAMES arginyl-trna synthetase
ORGANISM #formal_name Haemophilus influenzae
DATE 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 10-May-1996
ACCESSIONS A64131
REFERENCE A64000
#authors Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kierlavage, A.R.; Bult, C.J.; Tomb, J.F.; Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.; FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton, M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saudek, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.; Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
#journal Science (1995) 269:496-512
#title Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
#accession A64131
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-577 ##label TIGR
#cross-references GB:L42023; TIGR:HI1583
#note start codon was translated as Val
#note named as homolog to a protein from Escherichia coli
GENETICS
#gene args
CLASSIFICATION #superfamily arginine--trna ligase
KEYWORDS aminoacyl-trna synthetase; ligase; protein biosynthesis
SUMMARY #length 577 #molecular-weight 64732 #checksum 1459
Query Match 76.9%; Score 40; DB 6; Length 577;
Best Local Similarity 44.4%; Pred. No. 8.31e+01;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 246 lmvltetkd 254
|:||||:|
QY 1 LNVSLADTN 9
Search completed: Tue Jun 10 11:32:45 1997
Job time : 10 secs.
```

\*\*\*\*\*

W P E R E H (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligent Genetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 10 11:32:10 1997; MasPar time 1.87 Seconds  
Tabular output not generated. 101.859 Million cell updates/sec

Title: >US-08-231-565A-39  
Description: (1-9) from US08231565A.pap  
Perfect Score: 52  
Sequence: 1 LNVSLADTN 9

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 21.779; Variance 20.464; scale 1.064

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	52	100.0	668	PM17_HUMAN	PMEL 17 PROTEIN PRECU	2.07e-02
2	45	86.5	357	VINT_BPHKO	INTEGRASE.	1.59e+00
3	44	84.6	312	NULM_DROSU	NADH-UBIQUINONE OXIDO	2.86e+00
4	44	84.6	324	NULM_DROVA	NADH-UBIQUINONE OXIDO	2.86e+00
5	42	80.8	1046	MEXB_PSEAE	MULTIDRUG RESISTANCE	8.88e+00
6	42	80.8	2226	POLG_HPAV2	GENOME POLYPROTEIN (C	8.88e+00
7	42	80.8	2226	POLG_HPAV4	GENOME POLYPROTEIN (C	8.88e+00
8	42	80.8	2226	POLG_HPAV8	GENOME POLYPROTEIN (C	8.88e+00
9	42	80.8	2227	POLG_HPAVH	GENOME POLYPROTEIN (C	8.88e+00
10	42	80.8	2227	POLG_HPAVM	GENOME POLYPROTEIN (C	8.88e+00
11	42	80.8	2227	POLG_HPAVL	GENOME POLYPROTEIN (C	8.88e+00
12	41	78.8	420	AIAD_HUMAN	ALPHA-1-ANTITRYPSIN-R	1.54e+01
13	40	76.9	178	MYCB_RAT	B-MYC TRANSFORMING PR	2.64e+01
14	40	76.9	276	FDHD_MYCTU	FDHD PROTEIN HOMOLOG.	2.64e+01
15	40	76.9	290	YPUA_BACSU	HYPOTHETICAL 31.3 KD	2.64e+01
16	40	76.9	385	KUP1_CAEEL	KUP-1 PROTEIN	2.64e+01
17	40	76.9	577	SYR_HAEIN	ARGINYL-TRNA SYNTHETA	2.64e+01
18	40	76.9	873	YJ34_YEAST	HYPOTHETICAL 97.1 KD	2.64e+01
19	40	76.9	1034	ACRF_ECOLI	ACRIFLAVIN RESISTANCE	2.64e+01
20	40	76.9	1808	TENA_CHICK	TENASCIN PRECURSOR (T	2.64e+01
21	39	75.0	71	ARND_RAT	ARRESTIN-D (FRAGMENT)	4.47e+01
22	39	75.0	88	MINE_ECOLI	CELL DIVISION TOPOLOG	4.47e+01

23	39	75.0	407	6	LMPI_CRIGR	LYSOSOME-ASSOCIATED M	4.47e+01
24	39	75.0	433	9	TCR_STAAT	TETRACYCLINE RESISTAN	4.47e+01
25	39	75.0	458	9	TCR_BACST	TETRACYCLINE RESISTAN	4.47e+01
26	39	75.0	458	9	TCR_STRPN	TETRACYCLINE RESISTAN	4.47e+01
27	39	75.0	458	9	TCRB_BACSU	TETRACYCLINE RESISTAN	4.47e+01
28	39	75.0	458	9	TCR_STRAG	TETRACYCLINE RESISTAN	4.47e+01
29	39	75.0	458	9	TCR_STAHT	TETRACYCLINE RESISTAN	4.47e+01
30	39	75.0	459	9	TCR2_BACSU	TETRACYCLINE RESISTAN	4.47e+01
31	39	75.0	475	10	VG41_BPT4	PRIMASE-HELICASE (PRO	4.47e+01
32	39	75.0	484	7	PIGA_HUMAN	N-ACETYLGALUCOSAMINY	4.47e+01
33	39	75.0	569	4	GGTI_HUMAN	GAMMA-GLUTAMYLTRANSPE	4.47e+01
34	39	75.0	767	11	YNY7_YEAST	PUTATIVE CYSTEINYL-TR	4.47e+01
35	39	75.0	891	10	VP4A_VACCV	MAJOR CORE PROTEIN P4	4.47e+01
36	39	75.0	891	10	VP4A_VACCC	MAJOR CORE PROTEIN P4	4.47e+01
37	39	75.0	892	10	VP4A_VAVR	MAJOR CORE PROTEIN P4	4.47e+01
38	39	75.0	892	8	RAI6_SCHPO	DNA REPAIR PROTEIN RA	4.47e+01
39	39	75.0	944	7	NUFI_YEAST	NUFI PROTEIN (SPINDLE	4.47e+01
40	39	75.0	1389	9	TIM_DROME	TIMELESS PROTEIN.	4.47e+01
41	39	75.0	1402	1	ATCX_SCHPO	PROBABLE CALCIUM-TRAN	4.47e+01
42	39	75.0	2349	9	TPR_HUMAN	NUCLEOPROTEIN TPR.	4.47e+01
43	38	73.1	174	8	RLI6_SCHPO	PUTATIVE 60S RIBOSOMA	7.48e+01
44	38	73.1	440	3	ENO1_CANAL	ENOLASE 1 (EC 4.2.1.1	7.48e+01
45	38	73.1	482	5	LBP_RABIT	LIPOPOLYSACCHARIDE-BI	7.48e+01

ALIGNMENTS

RESULT	1	PM17_HUMAN	STANDARD;	PRT;	668 AA.
AC	P40967;				
DT	01-FEB-1995 (REL. 31, CREATED)				
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)				
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)				
DE	PMEL 17 PROTEIN PRECURSOR.				
GN	PMEL17.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92021023.				
RA	KWON B.S., CHINTANENI C., KOZAK C.A., COPELAND N.G.,				
RA	GILBERT D.J., JENKINS N., BARTON D., FRANCKE U., KOBAYASHI Y.,				
RA	KIM K.-K.				
RL	PROC. NATL. ACAD. SCI. U.S.A. 88:9228-9232(1991).				
CC	!- FUNCTION: COULD BE A MELANOGENIC ENZYME.				
CC	!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN MELANOCYTES.				
CC	!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).				
DR	EMBL; M77348; G190106; -				
DR	MIM; 155550; -				
KW	TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; MELANIN BIOSYNTHESIS; REPEAT.				
FT	SIGNAL	1	23		
FT	CHAIN	24	668		
FT	TRANSMEM	575	595		
FT	TRANSMEM	603	623		
FT	DOMAIN	217	307		
FT	DOMAIN	315	444		
FT	REPEAT	315	327		
FT	REPEAT	328	340		
FT	REPEAT	341	353		
FT	REPEAT	354	366		
FT	REPEAT	367	379		
FT	REPEAT	380	392		
FT	REPEAT	393	405		
FT	REPEAT	406	418		
FT	REPEAT	419	431		
FT	REPEAT	432	444		
FT	CARBOHYD	81	81		
FT	CARBOHYD	106	106		
FT	CARBOHYD	111	111		
FT	CARBOHYD	321	321		
FT	CARBOHYD	568	568		
FT	SEQUENCE	668 AA;	70992 MW;	6E8E1AF0 CRC32;	

```
Query Match      100.0%; Score 52; DB 7; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.07e-02;
Matches      9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 567 lnvsladtn 575
|||||
QY 1 LNVSLADTN 9

RESULT 2
ID VINT_BPHKO STANDARD; PRT; 357 AA.
AC P16407;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)
DE INTEGRASE.
GN INT.
OS BACTERIOPHAGE HK022.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; SIPHOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 83342457.
RA YAGIL E., DOLEV S., OBERTO J., KISLEV N., RAMAIAH N.,
RA WEISBERG R.A.;
RL J. MOL. BIOL. 207:695-717(1989).
CC -!- FUNCTION: INTEGRASE IS NECESSARY FOR INTEGRATION OF THE PHAGE INTO
CC THE HOST GENOME BY SITE-SPECIFIC RECOMBINATION. IN CONJUNCTION
CC WITH EXCISIONASE, INTEGRASE IS ALSO NECESSARY FOR EXCISION OF THE
CC PHORAGE FROM THE HOST GENOME.
CC -!- SIMILARITY: BELONGS TO THE "PHAGE" INTEGRASE FAMILY.
DR EMBL; X51962; G15760; -.
DR PIR; S04990; S04990.
KW DNA RECOMBINATION; DNA INTEGRATION.
FT ACT_SITE 342
FT DNA DURING STRAND CLEAVAGE AND REJOINING.
SQ SEQUENCE 357 AA; 40358 MW; B38FD29F CRC32;

Query Match      86.5%; Score 45; DB 10; Length 357;
Best Local Similarity 87.5%; Pred. No. 1.59e+00;
Matches      7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 251 lnlsladt 258
|||||
QY 1 LNVSLADT 8

RESULT 3
ID NULM_DROSU STANDARD; PRT; 312 AA.
AC P51937;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
GN ND1.
OS DROSOPHILA SUBOBSCURA (FRUIT FLY).
OG MITOCHONDRION.
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE OF 1-152 FROM N.A.
RX MEDLINE; 95106285.
RA BARRIO E., LATORRE A., MOYA A.;
RL J. MOL. EVOL. 39:478-488(1994).
RN [2]
RP SEQUENCE OF 50-312 FROM N.A.
RC STRAIN=SSP. TUE 3;
RA VOLZ-LINGENHIL A.;
RL SUBMITTED (JUN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
DR EMBL; U07288; G556934; -.
DR EMBL; X65130; G1070363; -.
DR FLYBASE; FBGN0012956; MT:ND1.
KW MITOCHONDRION; OXIDOREDUCTASE; NAD; TRANSMEMBRANE.
```

---

```
SQ SEQUENCE 312 AA; 35877 MW; 67FFD3D2 CRC32;

Query Match      84.6%; Score 44; DB 6; Length 312;
Best Local Similarity 66.7%; Pred. No. 2.86e+00;
Matches      6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 190 ltislaetn 198
|:|||||
QY 1 LNVSLADTN 9

RESULT 4
ID NULM_DROYA STANDARD; PRT; 324 AA.
AC P07710;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
GN ND1.
OS DROSOPHILA YAKUBA (FRUIT FLY).
OG MITOCHONDRION.
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2317.6 IVORY COAST;
RX MEDLINE; 86089137.
RA CLARY D.O., WOLSTENHOLME D.R.;
RL J. MOL. EVOL. 22:252-271(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84221393.
RA CLARY D.O., WAHLEITHNER J.A., WOLSTENHOLME D.R.;
RL NUCLEIC ACIDS RES. 12:3747-3762(1984).
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
DR EMBL; X03240; G12936; -.
DR PIR; D30020; D30020.
DR FLYBASE; FBGN0013183; MT:ND1.
DR PROSITE; PS00667; COMPLEX1_ND1_1.
DR PROSITE; PS00668; COMPLEX1_ND1_2.
KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION; TRANSMEMBRANE.
SQ SEQUENCE 324 AA; 37529 MW; 31D52170 CRC32;

Query Match      84.6%; Score 44; DB 6; Length 324;
Best Local Similarity 66.7%; Pred. No. 2.86e+00;
Matches      6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 190 ltislaetn 198
|:|||||
QY 1 LNVSLADTN 9

RESULT 5
ID MEXB_PSEAE STANDARD; PRT; 1046 AA.
AC P52002;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE MULTIDRUG RESISTANCE PROTEIN MEXB (MULTIDRUG-EFFLUX TRANSPORTER MEXB).
GN MEXB.
OS PSEUDOMONAS AERUGINOSA.
OC EUKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
OC PSEUDOMONADACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD10;
RX MEDLINE; 95058196.
RA POOLE K., HEINRICHS D.E., NESHAT S.;
RL MOL. MICROBIOL. 10:529-544(1993).
CC -!- FUNCTION: IMPLICATED IN THE SECRETION OF THE SIDEROPHORE
CC PYOVERDINE.
CC -!- INDUCTION: BY GROWTH UNDER SEVERE IRON LIMITATION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE ACRB/ACRF FAMILY.
```

```

DR EMBL: L11616; G438854; -.
KW TRANSMEMBRANE; INNER MEMBRANE; TRANSPORT.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 339 359 POTENTIAL.
FT TRANSMEM 370 390 POTENTIAL.
FT TRANSMEM 393 413 POTENTIAL.
FT TRANSMEM 441 461 POTENTIAL.
FT TRANSMEM 472 492 POTENTIAL.
FT TRANSMEM 539 559 POTENTIAL.
FT TRANSMEM 872 892 POTENTIAL.
FT TRANSMEM 894 914 POTENTIAL.
FT TRANSMEM 973 993 POTENTIAL.
FT TRANSMEM 1005 1025 POTENTIAL.
SQ SEQUENCE 1046 AA; 112778 MW; B682495C CRC32;

Query Match      80.8%; Score 42; DB 6; Length 1046;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 738 lgvsladln 746
| | | | | |
QY 1 LNVSLADTN 9

RESULT 6
ID POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DE 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN (COAT PROTEINS VP1 TO VP4; CORE PROTEINS P2A TO
DE P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)).
OS HEPATITIS A VIRUS (STRAIN 24A).
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; HEPATOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91162758.
RA LEMON S.M., MURPHY P.C., SHIELDS P.A., PING L.H., FEINSTONE S.M.,
RA CROMEANS T., JANSEN R.W.;
RL J. VIROL. 65:2056-2065(1991).
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC EMBL; M59810; G329588; -.
KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; RNA-DIRECTED RNA POLYMERASE;
KW HYDROLASE; THIOL PROTEASE.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; C67969A8 CRC32;

Query Match      80.8%; Score 42; DB 7; Length 2226;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lvnsladln 1375
| | | | | |
QY 1 LNVSLADTN 9

RESULT 7
ID POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DE 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN (COAT PROTEINS VP1 TO VP4; CORE PROTEINS P2A TO
DE P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)).
OS HEPATITIS A VIRUS (STRAIN 18F).
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; HEPATOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91162758.
RA LEMON S.M., MURPHY P.C., SHIELDS P.A., PING L.H., FEINSTONE S.M.,
RA CROMEANS T., JANSEN R.W.;
RL J. VIROL. 65:2056-2065(1991).
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC EMBL; M59808; G329586; -.
KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; RNA-DIRECTED RNA POLYMERASE;
KW HYDROLASE; THIOL PROTEASE.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
SQ SEQUENCE 2226 AA; 251152 MW; C67969A8 CRC32;

Query Match      80.8%; Score 42; DB 7; Length 2226;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lvnsladln 1375
| | | | | |
QY 1 LNVSLADTN 9

RESULT 7
ID POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DE 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN (COAT PROTEINS VP1 TO VP4; CORE PROTEINS P2A TO
DE P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)).
OS HEPATITIS A VIRUS (STRAIN 18F).
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; HEPATOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91162758.
RA LEMON S.M., MURPHY P.C., SHIELDS P.A., PING L.H., FEINSTONE S.M.,
RA CROMEANS T., JANSEN R.W.;
RL J. VIROL. 65:2056-2065(1991).
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC EMBL; M59808; G329586; -.
KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; RNA-DIRECTED RNA POLYMERASE;
KW HYDROLASE; THIOL PROTEASE.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
SQ SEQUENCE 2226 AA; 251152 MW; C67969A8 CRC32;

```

```

FT CHAIN          492 794 COAT PROTEIN VP1 (PID).
FT CHAIN          795 900 CORE PROTEIN P2A.
FT CHAIN          901 1087 CORE PROTEIN P2B.
FT CHAIN          1088 1422 CORE PROTEIN P2C.
FT CHAIN          1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN          1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN          1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN          1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE      2226 AA; 251292 MW; F50E6144 CRC32;

Query Match      80.8%; Score 42; DB 7; Length 2226;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lnnlaktin 1375
    |||:||||
QY 1 LNVSLADTN 9

RESULT 9
ID POLG.HPAVH STANDARD; PRT; 2227 AA.
AC P08617; P06443;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN (COAT PROTEINS VP1 TO VP4; CORE PROTEINS P2A TO
DE P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)).
OS HEPATITIS A VIRUS (STRAIN HM-175).
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; HEPATOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WILD TYPE;
RX MEDLINE; 87061253.
RA COHEN J.I., TICEHURST J.R., PURCELL R.H., BUCKLER-WHITE A.,
RA BAROUDY B.M.;
RL J. VIROL. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATTENUATED;
RX MEDLINE; 87175701.
RA COHEN J.I., ROSENBLUM B., TICEHURST J.R., DAEMER R.J., FEINSTONE S.M.,
RA PURCELL R.H.;
RL PROC. NATL. ACAD. SCI. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE; 85166289.
RA BAROUDY B.M., TICEHURST J.R., MIELE T.A., MAIZEL J.V. JR.,
RA PURCELL R.H., FEINSTONE S.M.;
RL PROC. NATL. ACAD. SCI. U.S.A. 82:2143-2147(1985).
CC -1- STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175
CC BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
DR EMBL; M14114; G329611; -
DR EMBL; M14707; G329583; -
DR EMBL; M16632; G329595; -
DR PIR; A25981; GNNYHM.
DR PIR; A25914; GNNYMK.
DR PIR; A03905; A03905.
KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; RNA-DIRECTED RNA POLYMERASE;
KW HYDROLASE; THIOL PROTEASE.
FT CHAIN          1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN          24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN          246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN          492 836 COAT PROTEIN VP1 (PID).
FT CHAIN          837 980 CORE PROTEIN P2A.
FT CHAIN          981 1087 CORE PROTEIN P2B.
FT CHAIN          1088 1422 CORE PROTEIN P2C.
FT CHAIN          1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN          1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN          1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN          1739 2227 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE      2227 AA; 251425 MW; 4AA67AF6 CRC32;

Query Match      80.8%; Score 42; DB 7; Length 2227;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

FT CHAIN          1088 1422 CORE PROTEIN P2C.
FT CHAIN          1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN          1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN          1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN          1739 2227 RNA-DIRECTED POLYMERASE 3D.
FT CHAIN          77 77 K -> R (IN ATTENUATED STRAIN).
FT VARIANT        764 764 E -> V (IN ATTENUATED STRAIN).
FT VARIANT        821 821 N -> S (IN ATTENUATED STRAIN).
FT VARIANT        821 821 A -> V (IN ATTENUATED STRAIN).
FT VARIANT        1052 1052 G -> A (IN ATTENUATED STRAIN).
FT VARIANT        1062 1062 K -> M (IN ATTENUATED STRAIN).
FT VARIANT        1118 1118 E -> K (IN ATTENUATED STRAIN).
FT VARIANT        1151 1151 F -> S (IN ATTENUATED STRAIN).
FT VARIANT        1163 1163 V -> I (IN ATTENUATED STRAIN).
FT VARIANT        1277 1277 H -> Y (IN ATTENUATED STRAIN).
FT VARIANT        1500 1500 D -> N (IN ATTENUATED STRAIN).
FT VARIANT        1805 1805 S -> T (IN ATTENUATED STRAIN).
FT VARIANT        1930 1930 S -> T (IN ATTENUATED STRAIN).
SQ SEQUENCE      2227 AA; 251506 MW; 4C2AD3BD CRC32;

```

```

Query Match      80.8%; Score 42; DB 7; Length 2227;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Db 1367 lnnlaktin 1375
    |||:||||
QY 1 LNVSLADTN 9

```

```

RESULT 10
ID POLG.HPAVH STANDARD; PRT; 2227 AA.
AC P13901;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN (COAT PROTEINS VP1 TO VP4; CORE PROTEINS P2A TO
DE P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)).
OS HEPATITIS A VIRUS (STRAIN MBB).
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; HEPATOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88045071.
RA PAUL A.V., TADA H., DER HELM K., WISSEL T., KIEHN R., WIMMER E.,
RA DEINHARDT F.;
RL VIRUS RES. 8:153-171(1987).
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT WHO HAD BEEN
CC INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
DR EMBL; M20273; G329607; -
DR PIR; J50303; GNNVHB.
KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; RNA-DIRECTED RNA POLYMERASE;
KW HYDROLASE; THIOL PROTEASE.
FT CHAIN          1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN          24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN          246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN          492 836 COAT PROTEIN VP1 (PID).
FT CHAIN          837 980 CORE PROTEIN P2A.
FT CHAIN          981 1087 CORE PROTEIN P2B.
FT CHAIN          1088 1422 CORE PROTEIN P2C.
FT CHAIN          1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN          1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN          1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN          1739 2227 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE      2227 AA; 251425 MW; 4AA67AF6 CRC32;

```

```

Query Match      80.8%; Score 42; DB 7; Length 2227;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```



```

Db 1367 lvnlnlakt 1375
|||:||||
QY 1 LNVSLADTN 9

RESULT 11
ID POLG.HPAVL STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE GENOME POLYPROTEIN (COAT PROTEINS VP1 TO VP4; CORE PROTEINS P2A TO
P2C; PROBABLE PROTEINS P3A TO P3C; RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)).
OS HEPATITIS A VIRUS (STRAIN LA).
NC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; PICORNAVIRIDAE; HEPATOVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85190549.
RA NAJARIAN R., CAPUT D., GEE W.W., POTTER S.J., RENARD A.,
RA MERRYWEATHER J., VAN NEST G., DINA D.;
RA PROC. NATL. ACAD. SCI. U.S.A. 82:2627-2631(1985).
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC EMBL; G02990; G329597; -
DR PIR; A03903; GNNVHR
KW POLYPROTEIN; COAT PROTEIN; CORE PROTEIN; RNA-DIRECTED RNA POLYMERASE;
KW HYDROLASE; THIOLESTER.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
FT CHAIN 2227 2227
SQ SEQUENCE 2227 AA; 251898 MW; 1A261852 CRC32;

Query Match 80.8%; Score 42; DB 7; Length 2227;
Best Local Similarity 77.8%; Pred. No. 8.88e+00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1367 lvnlnlakt 1375
|||:||||
QY 1 LNVSLADTN 9

RESULT 12
ID AJAU.HUMAN STANDARD; PRT; 420 AA.
AC P20848;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE ALPHA-1-ANTITRYPSIN-RELATED PROTEIN PRECURSOR.
PIL OR ATR.
OS HOMO SAPIENS (HUMAN).
NC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88314108.
RA BAO J.-J., REED-FOURQUET L., SIFERS R.N., KIDD V.J., WOO S.L.C.;
RA GENOMICS 2:165-173(1988).
CC -!- FUNCTION: COULD BE A SERINE PROTEASE INHIBITOR OR A PSEUDOGENE.
CC -!- SIMILARITY: WITH OTHER SERPIN PROTEASE INHIBITORS.
DR EMBL; M19684; G177814; -
DR EMBL; M19685; G177814; JOINED.

```

```

DR PIR; A38882; A28882.
DR HSP; P01009; 7API.
DR MIM; 107410; -.
DR PROSITE; PS00284; SERPIN.
KW SERPIN; SERINE PROTEASE INHIBITOR; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 ?
FT CHAIN ? 420
FT ACT_SITE 384 385
FT CARBOHYD 56 56
FT CARBOHYD 109 109
FT CARBOHYD 147 147
FT CARBOHYD 273 273
FT CARBOHYD 273 273
SQ SEQUENCE 420 AA; 47891 MW; BCBE4001 CRC32;

Query Match 78.8%; Score 41; DB 1; Length 420;
Best Local Similarity 62.5%; Pred. No. 1.54e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 106 lvnlnlakt 113
|||:||||
QY 1 LNVSLADTN 8

RESULT 13
ID MYCB.RAT STANDARD; PRT; 178 AA.
AC P15063;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE B-MYC TRANSFORMING PROTEIN (FRAGMENT).
GN MYCB.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=LIVER, AND BRAIN;
RX MEDLINE; 90082428.
RA ASKER C., STEINITZ M., ANDERSSON K., SUEMEGI J., KLEIN G.,
RA INGVARSSON S.;
RA INGVARSSON S.;
RL ONCOGENE 4:1523-1527(1989).
RN [2]
RX MEDLINE; 89096904.
RX SEQUENCE OF 1-120 FROM N.A.
RA INGVARSSON S., ASKER C., AXELSON H., KLEIN G., SUEMEGI J.;
RL MOL. CELL. BIOL. 8:3168-3174(1988).
DR EMBL; X17455; G55831; -.
DR EMBL; M21133; -; NOT_ANNOTATED_CDS.
DR PIR; A31198; TVRTBM.
DR PIR; A45502; A45502.
DR PROSITE; PS00038; HELIX_LOOP_HELIX.
DR TRANSFAC; T01563; -.
KW TRANSFORMING PROTEIN; NUCLEAR PROTEIN; DNA-BINDING.
FT NON_TER 178
FT NON_TER 178
SQ SEQUENCE 178 AA; 19463 MW; 8AEB4333 CRC32;

Query Match 76.9%; Score 40; DB 6; Length 178;
Best Local Similarity 66.7%; Pred. No. 2.64e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 3 lvnslangn 11
|||:||||
QY 1 LNVSLADTN 9

RESULT 14
ID FDHD.MCTU STANDARD; PRT; 276 AA.
AC Q10820;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE FDHD PROTEIN HOMOLOG.
GN MTCY274.30C.

```

```

OS MYCOBACTERIUM TUBERCULOSIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -!- SIMILARITY: TO OTHER BACTERIAL FDHD.
DR EMBL; 274024; E248774; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 276 AA; 29090 MW; BBF4E372 CRC32;

Query Match 76.9%; Score 40; DB 3; Length 276;
Best Local Similarity 66.7%; Pred. No. 2.64e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 245 lavsladas 253
| | | | | | | |
QY 1 LNVSLADTN 9

RESULT 15
ID YPUA_BACSU STANDARD; PRT; 290 AA.
AC P31847; P37951;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 31.3 KD PROTEIN IN LYSA-PPIB INTERGENIC REGION (ORFX19).
GN YPUA.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-I68 / MARGURG;
RX MEDLINE; 95020538.
RA SOROKIN A.V., ZUMSTEIN E., AZEVEDO V., EHRLICH S.D., SERROR P.;
RL MOL. MICROBIOL. 10:385-395(1993).
RN [2]
RP SEQUENCE OF 113-290 FROM N.A.
RX MEDLINE; 91345841.
RA YAMAMOTO J., SHIMIZU M., YAMANE K.;
RL AGRIC. BIOL. CHEM. 55:1615-1626(1991).
RN [3]
RP SEQUENCE OF 113-290 FROM N.A.
RX STRAIN-I68 / JH642;
RX MEDLINE; 94293776.
RA HERRLER M., BANG H., MARAHIEL M.A.;
RL MOL. MICROBIOL. 11:1073-1083(1994).
DR EMBL; L09228; G410118; -.
DR EMBL; D90189; G216392; -.
DR EMBL; S50865; -; NOT_ANNOTATED_CDS.
DR EMBL; X73898; G475053; -.
DR PIR; JU0473; JU0473.
DR PIR; S45536; S45536.
DR SUBTILIST; BG10511; YPUA.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 290 AA; 31294 MW; 6533D6F4 CRC32;

Query Match 76.9%; Score 40; DB 11; Length 290;
Best Local Similarity 62.5%; Pred. No. 2.64e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 218 lntvltids 225
| | | | | | | |
QY 1 LNVSLADT 8

Search completed: Tue Jun 10 11:32:18 1997
Job time : 8 secs.

```